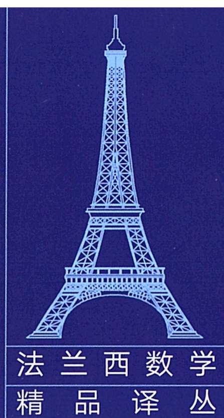


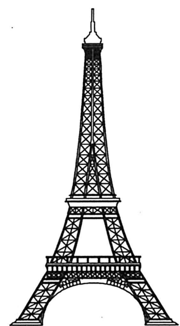
法兰西数学
精品译丛

线性与非线性泛函分析及其应用（上册 修订版）

Linear and Nonlinear Functional Analysis with Applications

Philippe G. Ciarlet 著
秦铁虎 童裕孙 译

高等教育出版社



法兰西数学
精品译丛

线性与非线性泛函分析及其应用（上册 修订版）

Linear and Nonlinear Functional Analysis with Applications

□ Philippe G. Ciarlet 著
□ 秦铁虎 童裕孙 译

高等教育出版社·北京

图字：01-2015-3396 号
Linear and Nonlinear Functional Analysis with Applications, by Philippe G. Ciarlet

图书在版编目（C I P）数据

线性与非线性泛函分析及其应用．上册 /（法）菲立普·G. 希阿雷（Philippe G.Ciarlet）著；秦铁虎，童裕孙译．-- 修订本．-- 北京：高等教育出版社，2020. 9

书名原文：Linear and Nonlinear Functional Analysis with Applications

ISBN 978-7-04-054552-4

Ⅰ．①线… Ⅱ．①菲… ②秦… ③童… Ⅲ．①泛函分析 Ⅳ．① O177

中国版本图书馆 CIP 数据核字（2020）第 120169 号

线性与非线性泛函分析及其应用
XIANXING YU FEIXIANXING FANHAN FENXI JIQI YINGYONG

策划编辑 吴晓丽　　责任编辑 吴晓丽　　封面设计 张　楠　　版式设计 童　丹
责任校对 吕红颖　　责任印制 存　怡

出版发行	高等教育出版社	网　　址	http://www.hep.edu.cn
社　　址	北京市西城区德外大街4号		http://www.hep.com.cn
邮政编码	100120	网上订购	http://www.hepmall.com.cn
印　　刷	北京佳顺印务有限公司		http://www.hepmall.com
开　　本	787mm×1092mm　1/16		http://www.hepmall.cn
印　　张	33.5	版　　次	2017 年 6 月第 1 版
字　　数	650 千字		2020 年 9 月第 2 版
购书热线	010-58581118	印　　次	2020 年 9月第 1 次印刷
咨询电话	400-810-0598	定　　价	89.00 元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换

物 料 号　54552-00

《法兰西数学精品译丛》序

随着解析几何及微积分的发明而兴起的现代数学, 在其发展过程中, 一批卓越的法国数学家发挥了杰出的作用, 做出了奠基性的贡献. 他们像灿烂的星斗散发着耀眼的光辉, 在现代数学史上占据着不可替代的地位, 在大学教科书、各种专著及种种数学史著作中都频繁地出现着他们的英名. 在他们当中, 包括笛卡儿、费马、帕斯卡、达朗贝尔、拉格朗日、蒙日、拉普拉斯、勒让德、傅里叶、泊松、柯西、刘维尔、伽罗瓦、庞加莱、嘉当、勒贝格、韦伊、勒雷、施瓦茨及利翁斯等这些耳熟能详的名字, 也包括一些现今仍然健在并继续做出重要贡献的著名数学家. 由于他们的出色成就和深远影响, 法国的数学不仅具有深厚的根基和领先的水平, 而且具有优秀的传统和独特的风格, 一直在国际数学界享有盛誉.

我国的现代数学, 在 20 世纪初通过学习西方及日本才开始起步, 并在艰难曲折中发展与成长, 终能在 2002 年成功地在北京举办了国际数学家大会, 在一个世纪的时间中基本上跟上了西方历经四个多世纪的现代数学发展的步伐, 实现了跨越式的发展. 这一巨大的成功, 根源于好几代数学家持续不断的艰苦奋斗, 根源于我们国家综合国力不断提高所提供的有力支撑, 根源于改革开放国策所带来的强大推动, 也根源于很多国际数学界同仁的长期鼓励、支持与帮助. 在这当中, 法兰西数学精品长期以来对我国数学界所起的积极影响, 法兰西数学的深厚根基、无比活力和优秀传统对我国数学家所起的不可低估的潜移默化作用, 无疑也是一个不容忽视的因素. 足以证明这一点的是: 在我国的数学家中, 有不少就曾经留学法国, 直接受到法国数学家的栽培和法兰西数学传统和风格的熏陶与感召, 而更多的人也或多或少地通过汲取法国数学精品的营养而逐步走向了自己的成熟与辉煌.

由于语言方面的障碍, 用法文出版的优秀数学著作在我国的传播受到了较大的限制. 根据一些数学工作者的建议, 并取得了部分法国著名数学家的热情支持, 高等教

育出版社决定出版《法兰西数学精品译丛》, 将法国的一些享有盛誉并有着重要作用与影响的数学经典以及颇具特色的大学与研究生数学教科书及教学参考书, 有选择地从法文原文分批翻译出版. 这一工作得到了国家自然科学基金委员会数学天元基金的支持和资助, 对帮助并推动我国读者更好地学习和了解法国的优秀数学传统和杰出数学成就, 进一步提升我国数学（包括纯粹数学与应用数学）的教学与研究工作的水平, 将是意义重大并影响深远的, 特为之序.

李大潜

2008 年 5 月

纪念我的父母 **Hélène** 和 **Gaston**

序　　言

在我们周围已经有很多优秀的教科书了, 为什么还要撰写另一部关于泛函分析及其应用的教科书呢?

除了把这样一种尝试视为作者个人兴趣的因素之外, 还有其他的原因: 一个原因是, 将线性及非线性泛函分析中最基本的定理收集在同一本书里, 这或许是撰写这部书的主要原动力; 另一个原因是, 在处理丰富的应用问题的同时也说明这些定理应用的广泛性.

在此书中讨论的关于对线性及非线性偏微分方程的应用包括: Korn 不等式及线性弹性的存在定理, 障碍问题, Babuška-Brezzi 上下确界条件, 流体力学中的 Stokes 和 Navier-Stokes 方程组的存在定理, 非线性弹性板中的 von Kármán 方程的存在定理, 以及非线性弹性中 John Ball 的存在性定理等. 各种各样的其他应用论题则选自数值分析及最优化理论. 例如, 逼近论, 多项式插值的误差估计, 数值线性代数, 最优化的基本算法, Newton 方法, 或有限差分法等.

我们也做了特别的努力, 以使本书更能满足教学上的要求. 其第 1 章实质上是对书中要用到的实分析及函数论中有关结果的复述. 而该章之后, 大部分定理都是自包含的, 给出了完整的证明[1]. 这些自包含的证明一般不太容易在其他文献中找到, 有些如果没有相关领域的扩展知识是很难得到的. 例如, 书中对于 Poincaré 引理, Laplace 算子的次椭圆性, Pfaff 方程组的存在定理, 或者曲面理论的基本定理等给出了这种自包含证明. 本书还包含诸多插图和 (约 400 道) 习题. 书中还给出了 (大部分是作为脚注) 有关史实的注记以及 (至少那些在有理由保证其真实性的前提下能追溯到的) 原始参考文献[2], 以对某些重要结果的产生提供一个原始思路.

---

[1] 定理左侧标记符号 $\flat$ 者, 表示其中不包含证明.

[2] 倾我所知来做这件事可能是一种冒险的尝试 ······

我相信, 本书覆盖了泛函分析中的大部分核心课题, 对线性及非线性应用感兴趣的分析学者在其职业生涯中都会接触到这些课题. 更具体地说, 线性泛函分析及其应用是第 2 章到第 6 章的主题, 而第 7 章到第 9 章的主题是非线性泛函分析及其应用.

当然, 为了能使本书的篇幅保持在一个合理限度内, 必须有所取舍. 一些更专门的课题, 如 Fourier 变换、小波、谱理论 (除紧自伴算子外) 以及与时间相关的偏微分方程等, 书中均未予以讨论.

在本科最后一年或研究生的水平上, 本书的内容可作为几个一学期课程的教科书, 例如, "线性泛函分析" "线性与非线性边值问题" "微分学及其应用" "微分几何导论" "非线性泛函分析" 以及 "数学弹性与流体力学" 等. 就此而言, 对于教师来说, 从内容目录中, 选取本书合适的部分作为这类课程的教科书是很容易的事. 实际上, 我非常愉快地讲授过这些课程. 最初是在巴黎第六大学 (Université Pierre et Marie Curie) 及香港城市大学, 后来也在奥斯汀的得克萨斯大学 (University of Texas at Austin), 康奈尔大学 (Cornell University), 复旦大学, 斯图加特大学 (University of Stuttgart), 苏黎世联邦理工大学 (ETH-Zürich) 以及苏黎世大学 (University of Zürich) 讲授过.

要求的主要预备知识是在一个合理的程度上知晓实分析, 即初等拓扑 (如连续性、紧性等), 距离空间的基本性质和 Lebesgue 积分, 以及单个或多个实变量的实值函数理论等. 为方便读者起见, 本书中需要用到的这些科目里的一些基本定义和定理都不加证明地收集在第 1 章中.

在撰写这部书期间, 我从 Liliana Gratie, George Dinca, Cristinel Mardare, Sorin Mardare, 以及 Pascal Azerad 等的评议中获益匪浅. 感谢他 (她) 们非常仔细地阅读了大部分章节, 并提出了许多有意义的改进意见. Bernard Dacorogna 与 Vicentiu Radulescu 也向我提供了宝贵的建议. 对他 (她) 们所有的人, 我表示衷心的感谢!

我还要感谢 Douglas N. Arnold, 他很早就对这一项目给予强有力的支持. 同时也要感谢 SIAM 编辑部的 Elizabeth Greenspan, Gina Rinelli 和 Lisa Briggeman, 与她们合作总是非常愉快的.

最后不可不提, 我要对我心目中的 "数学英才" 表示深切的感激和持久的敬意, 他们是 Laurent Schwartz, Richard S. Varga, Jacques-Louis Lions 和 Robert Dautray, 他们多年来的教诲与指导是无价可喻的.

我十分清楚本书肯定还存在一些不足之处. 例如, 可能所用的符号不一致, 无意中遗漏了应该标注的参考文献, 及引用的原始结果归属失当等. 但是, 任何探索 (数学或其他方面的), 即使主人公不太满意, 都得有个结局. 正如 Paul Halmos 在他的一篇论文[3] 的核心思想中, 以更恰当的方式所表述的, 任何数学家, 不管是纯粹还是应用数学家, 最好的办法是看一遍再复读一遍 (我理解他的意思): "大多数作者的最后一步是停笔, 但那是很艰难的一步."

---

[3] P. R. Halmos [1970]: How to write mathematics. *L'Enseignement Mathématique* **16**, 123–152.

这也就是为什么我预先表示欢迎所有的评论、注记、批评等的原因. 这些意见请发送到 mapgc@cityu.edu.hk, 它们或许可被采用于第二版中.

Philippe G. Ciarlet
2012 年 11 月于香港

目　录

第 2 章 赋范向量空间 43

第 3 章 Banach 空间 125

第 1 章　实分析和函数论: 快速回顾

引言

第 1 章作为对实分析的快速回顾, 按惯例包括集合论, 选择公理, 集合 $\mathbb{R}$ 和 $\mathbb{R}^n$ 的构造, 拓扑空间和距离空间的基本性质, 诸如连续性、紧性、完备性、连通性、单连通等概念; Tietze-Urysohn 扩张定理 (它在第 9 章多处有关键的应用); 还包括 $\mathbb{R}^n$ 上 Lebesgue 测度和 Lebesgue 积分的构造和主要性质: Radon-Nikodym 定理, Fatou 引理, Beppo Levi 单调收敛定理, Lebesgue 控制收敛定理, Tonelli 和 Fubini 定理, $\mathbb{R}^n$ 中的体积、面积和长度, 重积分的变量代换公式.

第 1 章还包括对多个实变量的实值函数理论若干方面的快速回顾. 特别是将介绍基本的函数空间, 如 $\mathcal{C}^m(\Omega)$ 和 $\mathcal{C}^m(\overline{\Omega})$, 其中 Ω 为 $\mathbb{R}^n$ 的开子集 (另一些函数空间, 如 Lebesgue 空间 $L^p(\Omega)$ 或 Sobolev 空间 $H^m(\Omega)$ 和 $W^{m,p}(\Omega)$ 将在以后各章作介绍). 在 $\mathbb{R}^n$ 的所有开子集中, 我们将单独介绍 $\mathbb{R}^n$ 中的域, 即开子集 $\Omega \subset \mathbb{R}^n$, 它是有界的、连通的、且边界有 Lipschitz 连续性, Ω 局部地在边界的同侧, 其原因是它保证了空间 $\mathcal{C}^1(\overline{\Omega})$ 中函数的基本 Green 公式 成立 ($\mathbb{R}^n$ 的域上的 Green 公式在后面将推广到 Sobolev 空间 $W^{m,p}(\Omega)$ 的函数上去, 见第 6 章).

此外, 我们假设读者已经熟悉有限维空间上的线性代数 (基、线性无关、矩阵、行列式等), 也熟悉多个实变量的实值函数的微分学的基本概念 (偏导数, Taylor 公式等).

这一章的目的实际上是以定理的形式给出本书中将要用到的一系列结果, 并列出为此将涉及的各种记号和定义.

考虑到读者比较熟悉这些结果, 本章中将略去相关的证明, 也不另行提供练习. 参考书将在文献注释中列出.

1.1　集合

最为普遍接受的集合理论是 Zermelo-Fraenkel 集合论. 它始于六个公理, 这里将不予罗列, 而仅仅叙述相应的推论.

这里将不再一一叙述诸如元素与集合等概念, 诸如 "$=, \neq, \in, \notin$" 等记号, 诸如 "导致" "对所有的" "存在" "满足" 等词或词组, 我们假设它们都是在直观或通常意义下可以理解的.

设 X 为一个集合. 记号 $A \subset X$ 或 $X \supset A$ 表示 A 是 X 的一个子集, 即当 $x \in A$ 时, $x \in X$. 记号 $A \subsetneqq X$ 或 $X \supsetneqq A$ 表示 A 是 X 的真子集, 即 $A \subset X$, 但 $A \neq X$.

设 X 为一个集合, 用 $\mathcal{P}(X)$ 表示以 X 的所有子集为元素全体的集合. 特别地, $\mathcal{P}(X)$ 包含空集 $\varnothing$ (由集合论公理, 它是存在的) 和集合 X 自身. 如果 $X \neq \varnothing$, $x \in X$, 则 X 的仅含元素 x 的子集记作 $\{x\}$.

设 X 为一个集合, $A \subset X$, 称 X 的由

$$X - A := \{x \in X;\ x \notin A\}$$

定义的子集为 A 关于 X 的余集, 如对集合 X 并无歧义, 也简称为 A 的余集.

如果 A 和 B 是集合 X 的子集, 其并和交分别定义为

$$A \cup B = \{x \in X;\ x \in A \text{ 或 } x \in B\},$$
$$A \cap B = \{x \in X;\ x \in A \text{ 且 } x \in B\}.$$

如果 $A \cap B = \varnothing$, 称集合 A 与 B 互不相交.

设 X 和 Y 为两个集合, 称由 $x \in X$ 和 $y \in Y$ 组成的序对 (x, y) 全体组成的集合

$$X \times Y := \{(x, y);\ x \in X,\ y \in Y\}$$

为 X 与 Y 的乘积.

集合 X 上的一个关系 $\mathcal{R}$ 是指乘积 $X \times X$ 的任一子集, 即 $\mathcal{R}$ 由一些特定的序对 (x, y) 组成, 其中 $x \in X$, $y \in Y$.

X 上的等价关系是指 X 上的满足下述条件的关系 $\mathcal{R}$, 其中 $(x, y) \in \mathcal{R}$, 记作 $x \sim y$, 满足:

自反性: 对所有 $x \in X$, 有 $x \sim x$,

对称性: 当 $x \sim y$ 时, $y \sim x$,

传递性: 当 $x \sim y$ 且 $y \sim z$ 时, $x \sim z$.

等价地, 这些条件即对任何 $x \in X$, 有 $(x, x) \in \mathcal{R}$; 如果 $(x, y) \in \mathcal{R}$, 则 $(y, x) \in \mathcal{R}$; 如果 $(x, y) \in \mathcal{R}$ 且 $(y, z) \in \mathcal{R}$, 则 $(x, z) \in \mathcal{R}$.

设 $\mathcal{R}$ 是集合 X 上的一个等价关系, $x \in X$, x 的模 $\mathcal{R}$ 的等价类定义为 X 的子集

$$\dot{x} := \{y \in X;\ y \sim x\}.$$

x 的元素的两个等价类或者相等或者互不相交.

商集 $X/\mathcal{R}$ 是指由 X 的元素模 $\mathcal{R}$ 的等价类全体组成的 $\mathcal{P}(X)$ 的子集.

上述所有的定义与性质仅依赖 Zermelo-Fraenkel 集合论的前五个公理. 第六个公理, 即无限性公理尤为重要. 因为由此一方面可导出由所有的自然数: $0, 1, 2, \cdots$ 组成的集合

$$\mathbb{N} = \{0, 1, 2, \cdots\}$$

的存在性, 同时, 也提供了用数学归纳法作证明的可能性: 为证明某性质对所有的 $n \in \mathbb{N}$ 成立, 只要证明它对 $n = 0$ 成立, 而且如果对 $n \in \mathbb{N}$ 成立, 则它对 $n+1$ 也成立. $\mathbb{N}$ 的一些特殊子集将用一些简明的记号表示, 如 $\{0, 1, 2, \cdots, n\}$, $\{j \in \mathbb{N};\ 1 \leqslant j \leqslant n\} = \{1, 2, \cdots, n\}$, $\{n \in \mathbb{N};\ n \geqslant n_0\} = \{n_0, n_0 + 1, \cdots\}$ 等.

1.2 映射

设 X 和 Y 为两个非空集合. X 到 Y 的映射或函数是指乘积 $X \times Y$ 的一个子集 f, 它满足对每个 $x \in X$, 存在唯一的 $y \in Y$ 使得 (x, y) 属于 f. 这里的元素 y 记作 $f(x)$ 或 y_x. 在记号 y_x 中, x 称为指标. 记号

$$f : X \to Y \quad \text{或} \quad X \xrightarrow{f} Y$$

表示 X 和 Y 是两个集合, f 是 X 到 Y 的映射. 带有显式表示 $f(x)$ 的记号

$$f : x \in X \to f(x) \in Y$$

用于定义映射 f.

设 X 是一个集合, 映射 $x \in X \to x \in X$ 称为 X 上的恒等映射, 记作 id 或 id_X, 当 X 为向量空间时, 也记作 I 或 I_X.

设 A 是集合 X 的一个子集, 定义函数 $\chi_A : X \to \mathbb{R}$ 为

$$\chi_A(x) := \begin{cases} 1, & x \in A, \\ 0, & x \notin A, \end{cases}$$

并称之为 A 的特征函数.

设 $f : X \to Y$ 为一个映射, 所谓 X 的子集 A 在 f 下的直接像是指 Y 的子集 $f(A)$, 其定义为

$$f(A) := \{y \in Y;\ \text{存在 } x \in A,\ \text{使得 } y = f(x)\}.$$

Y 的子集 B 在 f 下的逆像定义为 X 的子集

$$f^{-1}(B) := \{x \in X;\ f(x) \in B\}.$$

设 $b \in Y$, 逆像 $f^{-1}(\{b\})$ 常简洁地用 (并不恰当但却方便的) 记号 $f^{-1}(b)$ 表示.

在使用记号 $f(A)$ 和 $f^{-1}(B)$ 时必须注意到: 记号 f 表示 X 到 Y 的映射, 而并非 $\mathcal{P}(X)$ 到 $\mathcal{P}(Y)$ 的映射 (虽似记号 $f(A)$ 所 "启示"). 同样地, 当 f 的逆映射存在时 (见下文), 记号 f^{-1} 表示 Y 到 X 的映射, 而并非 $\mathcal{P}(Y)$ 到 $\mathcal{P}(X)$ 的映射 (虽似记号 $f^{-1}(B)$ 所 "启示").

逆像 "保持所有的集合运算", 即满足

$$\begin{gathered}
\text{若 } B \subset \widetilde{B}, \text{ 则 } f^{-1}(B) \subset f^{-1}(\widetilde{B}),\\
f^{-1}(B \cup \widetilde{B}) = f^{-1}(B) \cup f^{-1}(\widetilde{B}),\\
f^{-1}(B \cap \widetilde{B}) = f^{-1}(B) \cap f^{-1}(\widetilde{B}),\\
f^{-1}(Y - B) = X - f^{-1}(B).
\end{gathered}$$

相反地, 直接像仅满足

$$\begin{gathered}
\text{若 } A \subset \widetilde{A}, \text{ 则 } f(A) \subset f(\widetilde{A}),\\
f(A \cup \widetilde{A}) = f(A) \cup f(\widetilde{A}),\\
f(A \cap \widetilde{A}) \subset f(A) \cap f(\widetilde{A}).
\end{gathered}$$

设有映射 $f: X \to Y$, 如果对每个 $y \in Y$, 至少存在一个 $x \in X$, 使得 $y = f(x)$, 则称 f 为满的或到 Y 上的, 也称作满射.

设有映射 $f: X \to Y$, 如果对每个 $y \in Y$, 至多存在一个 $x \in X$, 使得 $y = f(x)$, 则称 f 为单的或一对一的, 也称作单射. 如果 X 是 Y 的子集, 定义映射 $\iota: X \to Y$ 为对每个 $x \in X$, $\iota(x) = x$, 称之为由 X 到 Y 的典则单射.

设映射 $f: X \to Y$ 既是满的又是单的, 则称 f 为一个双射或一对一且到上的映射. 此时, 对每个 $y \in Y$, 存在唯一的 $x \in X$, 使得 $y = f(x)$, 由此定义的映射 $f^{-1}: y \in Y \to x \in X$ 称为 f 的逆映射.

设 $f: X \to Y$ 为一个映射, A 是 X 的一个子集. 对每个 $x \in A$, 令 $f|_A(x) := f(x)$, 这就定义了一个映射 $f|_A: A \to Y$, 称之为 f 在 A 上的限制.

设 $g: A \to Y$ 为一个映射, 其中 A 为 X 的一个子集, 如果映射 $f: X \to Y$ 满足 $f|_A = g$, 则称 f 为 g 的一个延拓.

设 $f: X \to Y$, $g: Y \to Z$ 为两个映射, 映射 $h: X \to Z$ 定义为对每个 $x \in X$, $h(x) = g(f(x))$, 称 h 为复合映射, 记作 $h = g \circ f$ 或 $h = gf$.

设 $f: X \times Y \to Z$ 为一个映射, a 为 X 中的一个元素, 定义映射 $f(a, \cdot): Y \to Z$ 为

$$f(a, \cdot): y \in Y \to f(a, y) \in Z,$$

称之为一个部分映射. 对给定的元素 $b\in Y$, 同样可定义部分映射 $f(\cdot,b):X\to Z$.

给定映射 $f:(x,y)\in X\times Y\to f(x,y)\in Z$, 分别称元素 $x\in X$ 和 $y\in Y$ 为 f 的第一个变元和第二个变元.

1.3 选择公理和 Zorn 引理

设 I 和 X 为两个非空集合. 所谓以 I 为指标集的 X 的一个元素族即一个映射 $f:I\to X$, 定义为 $f:i\in I\to x_i\in X$, 即视 I 中的元素为指标. 这样的元素族记作

$$(x_i)_{i\in I},$$

在集合 I 无歧义时也简记作 (x_i). 自然, 必须仔细区分 X 的元素族 $(x_i)_{i\in I}$ 和 X 的子集 $\cup_{i\in I}\{x_i\}$ (例如, 当所有的 $i\in I$ 均有 $x_i=a$ 时, 后者即由单个点 $a\in X$ 组成).

元素族 $(x_i)_{i\in I}$ 的子族 $(x_i)_{i\in J}$ 是指映射 $g:J\to X$, 满足 $J\subset I$, 且 $f|_J=g$.

如果 $I=\{1,\cdots,n\}$, 其中 $n\geqslant 1$, 称元素族 $(x_i)_{i\in I}$ 为一 n 维元, 记作

$$(x_j)_{j=1}^n \quad 或 \quad (x_1,\cdots,x_n).$$

如果 $I=\mathbb{N}$, 称元素族 $(x_i)_{i\in I}$ 为一序列, 记作

$$(x_n)_{n=0}^\infty \text{ 或 } (x_0,x_1,\cdots,x_n,\cdots) \text{ 或 } (x_n)_{n\geqslant 0} \text{ 或简记为 } (x_n).$$

还有一些意义自明的记号, 如

$$(x_n)_{n\geqslant 0},\ 如果\ I=\{n_0,n_0+1,\cdots\},\ (x_n)_{n=n_0}^\infty \text{ 或 } (x_n)_{n\geqslant n_0} \text{ 等}.$$

一个序列的子列是指自身也是一个序列的子族. 例如, 给定一个严格单调增加映射 $\sigma:\mathbb{N}\to\mathbb{N}$ (即对任何 $n\in\mathbb{N}$, 有 $\sigma(n)<\sigma(n+1)$), 序列 $(x_{\sigma(n)})_{n=0}^\infty$ 是序列 $(x_n)_{n=0}^\infty$ 的一个子序列. 下文将经常使用这个记号表示子序列.

设 I 和 X 为两个集合. 所谓以 I 为指标集的 X 的子集族 $(A_i)_{i\in I}$ 是指一族 $i\in I\to A_i\in\mathcal{P}(X)$ (即出现在元素族定义中的映射, 它取值于 $\mathcal{P}(X)$ 而非 X).

给定集合 X 的一个子集族 $(A_i)_{i\in I}$, 并集 $\cup_{i\in I}A_i$ 和交集 $\cap_{i\in I}A_i$ 分别定义为

$$\bigcup_{i\in I}A_i=\{x\in X;\ 存在\ i\in I,\ 使得\ x\in A_i\},$$

$$\bigcap_{i\in I}A_i=\{x\in X;\ 对所有\ i\in I,\ 均有\ x\in A_i\}.$$

另外还有一些意义自明的记号, 如当 $I=\{0,1,\cdots,n\}$ 时的

$$\bigcup_{i=0}^n A_i,\quad \bigcap_{i=0}^n A_i$$

和当 $I=\mathbb{N}$ 时的

$$\bigcup_{i=0}^{\infty} A_i, \quad \bigcap_{i=0}^{\infty} A_i.$$

给定集合 X 的一族子集 $(A_i)_{i\in I}$, 定义其不相交的并 $\sqcup_{i\in I}A_i$ 为

$$\bigsqcup_{i\in I} A_i := \bigcup_{i\in I}\{(x,i); x\in A_i\}.$$

即不相交的并 $\sqcup_{i\in I}A_i$ 是乘积集合 $(\cup_{i\in I}A_i)\times I$ 的子集, 而该乘积集合又是乘积集合 $X\times I$ 的一个子集.

如果 A,B 是 X 的两个子集, 并集 $A\cup B$ 和交集 $A\cap B$ 即当取 $I:=\{1,2\}$, $A_1:=A$, $A_2:=B$ 时的并集 $\cup_{i\in I}A_i$ 和交集 $\cap_{i\in I}A_i$.

下面列出一些常用的等式:

$$A\cup\left(\bigcap_{i\in I} A_i\right)=\bigcap_{i\in I}(A\cup A_i), \quad A\cap\left(\bigcup_{i\in I} A_i\right)=\bigcup_{i\in I}(A\cap A_i),$$
$$A-\bigcup_{i\in I} A_i=\bigcap_{i\in I}(X-A_i), \quad X-\bigcap_{i\in I} A_i=\bigcup_{i\in I}(X-A_i),$$

后面两个等式就是 de Morgan 定律.

设 $(A_i)_{i\in I}$ 是集合 X 的一族子集, 其中 $I\neq\varnothing$. 乘积 $\prod_{i\in I}A_i$ 定义为所有映射 $f: I\to X$ 的集合, 其中 f 满足对每个 $i\in I$, $f(i)\in A_i$. 正因为对每个 $i\in I$, 它 “选取” 了一个元素 $f(i)\in A_i$, 所以这种映射也称为*选择函数*.

并集 $\cup_{i\in I}A_i$ 和交集 $\cap_{i\in I}A_i$ 作为 X 的子集, 其定义并无特殊的困难. 而乘积 $\prod_{i\in I}A_i$ 的定义直接列出了一个问题, 即并未保证至少有一个选择函数 f 的存在性, 这就是 Ernest Zermelo 在 1904 年要提出下述选择公理的原因.

选择公理 设 $(A_i)_{i\in I}$ 为某集合的一族子集, 如果 $I\neq\varnothing$, 且对任何 $i\in I$, $A_i\neq\varnothing$, 则 $\prod_{i\in I}A_i\neq\varnothing$.

1963 年, P. J. Cohen 在一篇里程碑式的文章[1] 里证明了选择公理独立于 Zermelo-Fraenkel 集合论的六个公理.

对于乘积 $\prod_{i\in I}A_i$ 还有下列常用记号:

当 $I=\{1,2\}$ 时, $A_1\times A_2$ (见 1.1 节),

当 $I=\{1,2,\cdots,n\}$ 时, $\displaystyle\prod_{i=1}^{n}A_i$,

当 $I=\{1,2,\cdots,n\}$ 且对所有 $1\leqslant i\leqslant n$, $A_i=A$ 时, A^n.

[1] P. J. Cohen [1963]: The independence of the continuum hypothesis. Proceedings of the National Academy of Sciences, USA 50, 1143—1148.

$\prod_{i\in I} A_i$ 的元素作为一个选择函数 f 也记作 $x=(x_i)_{i\in I}$, 其中 $x_i := f(i)$, 因而每个 $i\in I$ 被视为一个足标 (见 1.2 节). 对每个 $i\in I$, 称 $x_i \in A_i$ 为 x 的第 i 个坐标. 这个记号与乘积 $\mathbb{K}^n$ 中元素 $(x_i)_{i=1}^n$ 或 $\prod_{i=0}^{\infty} A_i$, 其中 $A_i=\mathbb{K}$ $(i\in\mathbb{N})$ 中元素 $(x_i)_{i=0}^{\infty}$ 的记号相符.

实际上, 根据问题的困难所在, 选择公理常以其等价定理形式使用. Zorn 引理 (下文定理 1.3-1) 就是一个例子. 注意, 虽然选择公理的叙述较为直观, Zorn 引理则并非如此.

为了叙述这个引理, 需要几个定义.

设 X 为一个集合, $\mathcal{R}$ 为 X 上的一个关系. 如果 $(x,y)\in\mathcal{R}$ 时记作 $x\preccurlyeq y$, 满足

自反性: 对所有的 $x\in X$, 有 $x\preccurlyeq x$,

反称性: 当 $x\preccurlyeq y$ 且 $y\preccurlyeq x$ 时 $x=y$,

传递性: 当 $x\preccurlyeq y$ 且 $y\preccurlyeq z$ 时 $x\preccurlyeq z$,

则称 X 为半序集, $\mathcal{R}$ 是 X 上的一个半序关系. 等价地, 上述三条性质即对所有 $x\in X$, 有 $(x,x)\in\mathcal{R}$; 当 $(x,y)\in\mathcal{R}$ 且 $(y,x)\in\mathcal{R}$ 时 $x=y$; 当 $(x,y)\in\mathcal{R}$, $(y,z)\in\mathcal{R}$ 时 $(x,z)\in\mathcal{R}$.

例如, 在 $\mathbb{R}^n$ 中由 "对所有 $1\leqslant i\leqslant n$, 均有 $x_i\leqslant y_i$ 规定 $x=(x_i)_{i=1}^n \preccurlyeq y=(y_i)_{i=1}^n$" 定义了一个半序; 在由 X 的子集全体组成的集合 $\mathcal{P}(X)$ 上由 "$A\subset B$ 规定 $A\preccurlyeq B$" 定义了一个半序.

记号 $y\succcurlyeq x$ 表示 $x\preccurlyeq y$. 记号 $x\prec y$ 或 $y\succ x$ 表示 $x\preccurlyeq y$ 且 $x\neq y$.

设 A 是半序集 X 的一个子集, 如果任何两个元素 $a\in A$ 和 $b\in A$ 均是可比较的, 即或者 $a\preccurlyeq b$, 或者 $b\preccurlyeq a$ (如果 $a\preccurlyeq b$ 与 $b\preccurlyeq a$ 同时成立, 则 $a=b$), 就称 A 为全序的. 如果全序集 A 是有限集, 即形如 $A=\cup_{i=1}^m\{a_i\}$, 则存在 $1\leqslant i_0\leqslant m$, 使得 $a_i\preccurlyeq a_{i_0}$ 对一切 $1\leqslant i\leqslant m$ 成立. 例如, 如果 $\mathcal{P}(X)$ 的有限子集 $\{A_1,A_2,\cdots,A_m\}$ 对于以包含关系为序关系而言是全序的, 则存在 $1\leqslant m_0\leqslant m$, 使得对所有的 $1\leqslant i\leqslant m$, 有 $A_i\subset A_{m_0}$, 从而 $A_{m_0}=\cup_{i=1}^m A_i$, 这个结论将常被用到.

设 A 是半序集 X 的一个子集. 如果 $b\in X$ 满足对一切 $a\in A$ 均有 $a\preccurlyeq b$, 则称 b 为 A 的一个上界. 注意, 此时 A 中每个元素均与 b 可比较, 但 b 未必属于 A.

设 X 为一半序集. 如果 $m\in X$, 满足对 X 中每个可与 m 比较的元素 x 均有 $x\preccurlyeq m$, 则称 m 为极大元; 这又等价于如果 $x\in X$ 满足 $m\preccurlyeq x$, 必有 $x=m$. 注意这样的 m 未必与每个 $x\in X$ 可比较.

下面的结果等价于选择公理.

定理 1.3-1 (Zorn 引理) 设 X 为非空半序集, 它的每个全序子集在 X 中均有上界, 则 X 至少有一个极大元.

Zorn 引理在导出某些数学对象存在性时是非常有用的工具. 例如, 可以用以证

明 $\mathbb{R}$ 上 Lebesgue 不可测集的存在性 (1.14 节); 证明任何向量空间具有 Hamel 基 (定理 2.1-1); 证明任何向量空间均可赋范 (定理 2.2-8); 证明任何内积空间具有极大的规范正交系 (定理 4.8-4); 证明基本的 Hahn-Banach 定理, 由此又可导出线性泛函延拓的存在性 (定理 5.8-1 和定理 5.9-1).

注意, 每次应用 Zorn 引理时, 均须验证相应的集合 X 是非空的.

1.4 集合 $\mathbb{R}$ 和 $\mathbb{C}$ 的构造

由无限性公理 (第 1.1 节) 导致其存在的集合 $\mathbb{N}$, 可以构造所有整数的集合

$$\mathbb{Z} := \{\cdots, -2, -1, 0, 1, 2, \cdots\},$$

它作为商集 $(\mathbb{N}\times\mathbb{N})/\mathcal{R}$, 其中 $\mathbb{N}\times\mathbb{N}$ 上的等价关系 $\mathcal{R}$ 定义为

$$((m,n),(m',n')) \in \mathcal{R} \text{ 当且仅当 } m+n' = m'+n.$$

带有加法和乘法的集合 $\mathbb{Z}$ 成为一个交换环, 也是一个整环 (即若 $mp = mq$ 且 $m \neq 0$, 则 $p = q$), 按 $\leqslant$ 它是全序集 (全序的定义见 1.3 节).

由集合 $\mathbb{Z}$ 可构造所有有理数的集合 $\mathbb{Q}$, 即 $\mathbb{Z}\times(\mathbb{Z}-\{0\})$ 的按下列等价关系的等价类全体的集合, 这里 $(m,n)\sim(p,q)$ 当且仅当 $mq = np$. 带有运算 $+$ 和 $\times$, 集合 $\mathbb{Q}$ 是一个全序交换域. 域 $\mathbb{Q}$ 具有 Archimedes 性, 即给定有理数 $r > 0$ 和 $s > 0$, 则存在整数 $n \geqslant 0$, 使得 $nr > s$. $r \in \mathbb{Q}$ 的绝对值定义为当 $r \geqslant 0$ 时, $|r| := r$; 当 $r < 0$ 时, $|r| := -r$.

称有理数序列 $(r_n)_{n=1}^{\infty}$ 为 Cauchy 序列, 是指对任何 $\varepsilon > 0$, 存在整数 $m_0 = m_0(\varepsilon) \geqslant 1$, 使得当所有的 m, $n \geqslant m_0$ 时, $|r_m - r_n| \leqslant \varepsilon$. 由集合 $\mathbb{Q}$ 可以构造所有实数的集合 $\mathbb{R}$, 它是由有理数的 Cauchy 序列全体组成的集合, 模下述等价关系的一切等价类的集合, 这里 $((r_n)_{n=1}^{\infty},(s_n)_{n=1}^{\infty}) \in \mathcal{R}$ 当且仅当对任意给定的 $\varepsilon > 0$, 存在整数 $n_0 = n_0(\varepsilon) \geqslant 1$, 对所有的 $n \geqslant n_0$, $|r_n - s_n| \leqslant \varepsilon$. 赋以运算 $+$ 和 $\times$ 及全序关系 $\leqslant$, 集合 $\mathbb{R}$ 是一个全序的 Archimedes 交换域. $x \in \mathbb{R}$ 的绝对值类似地定义为当 $x \geqslant 0$ 时 $|x| := x$, 当 $x < 0$ 时 $|x| := -x$. 另外, 集合 $\mathbb{R}$ 上也可用 Dedekind 分割的方法构造.

扩充的实数系是指由集合 $\mathbb{R}$ 添上另两个分别记作 $-\infty$ 和 ∞ 的元素组成的集合 $\{-\infty\}\cup\mathbb{R}\cup\{\infty\}$, 它满足一些通常的规则. 例如, 对 $x \in \mathbb{R}$, $-\infty < x$, $x + \infty = \infty$ 等, 但 $-\infty + \infty$ 则无定义.

最后, 复数全体组成的交换域 $\mathbb{C}$ 可按通常方法由集合 $\mathbb{R}$ 构造. 如果 $z \in \mathbb{C}$, 用 $\mathrm{Re}\, z$ 和 $\mathrm{Im}\, z$ 分别表示 z 的实部和虚部, 即 $z = \mathrm{Re}\, z + i\mathrm{Im}\, z$. $z \in \mathbb{C}$ 的绝对值定义为 $|z| := \sqrt{|\mathrm{Re}\, z|^2 + |\mathrm{Im}\, z|^2}$.

通常为方便起见, 用同一个字母 $\mathbb{K}$ 既表示域 $\mathbb{R}$, 也表示域 $\mathbb{C}$, 此时, 称 $\mathbb{K}$ 的元素为标量.

如上所述构造了集合 $\mathbb{R}$ 和 $\mathbb{C}$, 就可以建立起集合 $\mathbb{R}$ 和 $\mathbb{C}$ 的各种性质. 下面的定理汇集了其中最重要的一些性质. 实或复数 Cauchy 序列与有理 Cauchy 序列定义类似.

定理 1.4-1 (a) 集合 $\mathbb{R}$ 和 $\mathbb{C}$ 是完备的, 即实或复数的任何 Cauchy 序列 $(x_n)_{n=1}^{\infty}$ 收敛于一个实数或相应的一个复数, 即存在 $x \in \mathbb{R}$ 或相应的 $x \in \mathbb{C}$, 对任意给定的 $\varepsilon > 0$, 存在整数 $n_0 = n_0(\varepsilon) \geqslant 1$, 对所有的 $n \geqslant n_0$, 有

$$|x_n - x| \leqslant \varepsilon.$$

(b) $\mathbb{R}$ 和 $\mathbb{C}$ 的 Bolzano-Weierstrass 性质: 任何有界的实数列或复数列 $(x_n)_{n=1}^{\infty}$, 即存在 $M \in \mathbb{R}$, 对所有的 $n \geqslant 1$, 均有 $|x_n| \leqslant M$, 必定包含收敛子列.

(c) 设 A 是 $\mathbb{R}$ 的非空子集, 且在 $\mathbb{R}$ 中有上界 (1.3 节). 则存在 $a \in \mathbb{R}$ 作为 A 的最小上界, 或上确界; 即 a 是 A 的上界, 且对 A 的任何上界 $b \in \mathbb{R}$, 均有 $a \leqslant b$.

类似地, $\mathbb{R}$ 的非空子集如在 $\mathbb{R}$ 中有下界, 则必定在 $\mathbb{R}$ 中有最大下界, 即下确界.

当实数列或复数列 $(x_n)_{n=1}^{\infty}$ 收敛于 x 时 (如 (a) 所定义的) 也记作

$$x = \lim_{n\to\infty} x_n, \quad 或\ x_n \xrightarrow[n\to\infty]{} x, \quad 或当\ n \to \infty\ 时\ x_n \to x.$$

集合 X 到 $\mathbb{R}$ 或 $\mathbb{C}$ 的映射称之为相应的实值或复值函数.

集合 X 到 $\mathbb{R}^n$ 或所有 $m \times n$ 矩阵集合的映射称为向量域或矩阵域.

1.5 基数; 有限集和无限集

设 X 为一个集合. 易见由 "存在 A 到 B 的一个双射", 其中 A, B 为 X 的子集, 在集合 $\mathcal{P}(X)$ 上定义了一个等价关系 (1.1 节). 商集 $\mathcal{P}(X)/\mathcal{R} \subset \mathcal{P}(\mathcal{P}(X))$ 的元素称为集合 X 的子集的基数. 如果 A 为 X 的一个子集, 则 A 的模 $\mathcal{R}$ 的等价类, 即集合 $\mathcal{P}(\mathcal{P}(X))$ 的相应元素, 就是 A 的基数, 记作

$$\text{card}\ A.$$

值得注意的是由于下述基本定理, 集合 $\mathcal{P}(X)/\mathcal{R}$ 是可全序的 (全序定义见 1.3 节).

定理 1.5-1 集合 X 的子集全体的基数的集合 $\mathcal{P}(X)/\mathcal{R}$ 按关系 $\boldsymbol{R}$ 是全序的, 这里 $(\text{card}\ A, \text{card}\ B) \in \boldsymbol{R}$ 意即存在 A 到 B 的一个单射. 等价地, $(\text{card}\ A, \text{card}\ B) \in \boldsymbol{R}$ 当且仅当存在 B 到 A 上的一个满射.

下面对这个结果的证明给予提示 (粗看上去证明似乎无实质意义, 实际上却并非平凡). 第一, $\boldsymbol{R}$ 的定义并无歧义. 这是因为如果 $\text{card}\ A = \text{card}\ \widetilde{A}$, $\text{card}\ B = \text{card}\ \widetilde{B}$, 且存在 A 到 B 的单射, 则必存在 $\widetilde{A}$ 到 $\widetilde{B}$ 的单射.

第二, 必须证明 $\boldsymbol{R}$ 是集合 $\mathcal{P}(X)$ 上的一个半序. 关于自反性和传递性可以直接验证, 反称性则不然. 为此必须证明如果存在 A 到 B 的单射和 B 到 A 的单射, 则必存在 A 到 B 上的一个双射[2).

第三, 必须证明对于关系 $\boldsymbol{R}$, $\mathcal{P}(X)/\mathcal{R}$ 是全序集, 即给定任何两个子集 $A \subset X$ 和 $B \subset X$, 或者存在 A 到 B 的单射, 或者存在 B 到 A 的单射, 这两个性质互不相斥. 这一部分的证明[3), 要用到选择公理.

最后, 还必须证明: 存在 A 到 B 的单射的充要条件是存在 B 到 A 的双射, 其中充分性的证明也要用到选择公理.

如同 1.3 节, 后面将使用更为简明的记号

$$\text{card}\, A \preccurlyeq \text{card}\, B \quad \text{或} \quad \text{card}\, A \prec \text{card}\, B$$

分别表示 $(\text{card}\, A,\ \text{card}\, B) \in \boldsymbol{R}$ 或 $(\text{card}\, A,\ \text{card}\, B) \in \boldsymbol{R}$ 且 $\text{card}\, A \neq \text{card}\, B$.

现在我们来比较集合的基数. 即便作为某集合的子集, 尚未给出它们的基数, 然而, 在 Zermelo-Fraenkel 集合论中 (1.1 节), 这些集合的并总是确定的, 所以这种比较并没有困难.

下述的结果[4) 可粗略地表述为: 不存在 “最大” 的基数 (对于关系 $\boldsymbol{R}$).

定理 1.5-2 *设 X 为任一集合, 则不存在 X 到集合 $\mathcal{P}(X)$ 上的双射, 即关系*

$$\text{card}\, X \prec \text{card}\, \mathcal{P}(X)$$

成立.

如果 $X = \varnothing$ 或者存在整数 $n \geqslant 1$, 使得 $\text{card}\, X = \text{card}\,\{1, \cdots, n\} = n$, 则称 X 为*有限集*. 非有限的集合称为*无限集*. 特别地, 如果 $\text{card}\, X = \text{card}\, \mathbb{N}$, 则称 X 为*可数无限的*, 若既非有限, 亦非可数无限, 则称集合为*不可数无限*. 注意, 有些作者把有限集和可数无限集均称为可数集.

下面的定理列出了无限集的一些重要性质. (a) 和 (b) 的证明都需要用选择公理. 性质 (a) 说明 $\text{card}\, \mathbb{N}$ 是最小的无限基数.

定理 1.5-3 (a) *设 X 为无限集, 则*

$$\text{card}\, \mathbb{N} \preccurlyeq \text{card}\, X.$$

---

2) 这就是由 Felix Bernstein (1878—1956) 在 1897 年证明的著名定理的内容.

3) E. Zermelo [1904]: Beweis dass jede Menge wohlgeordnet werden kann. Mathematische Annalen **LIX**, 514–516.

4) 基数理论当归功于 Georg Cantor (1845—1918), 他在其有深远影响 (但也在很长时期中引起包括一些著名数学家激烈争辩的) 著作中阐述了这一理论. G. Cantor [1899]: Beiträge zur Begründung der transfiniten Mengenlehre. Georg Olms Verlag (English translation: Contributions to the Founding of Transfinite Numbers, Dover, New York, 1955).

(b) 设 X 为无限集, 则

$$\text{card}\,(X \times X) = \text{card}\,X.$$

(c) 实数集 $\mathbb{R}$ 的基数满足

$$\text{card}\,\mathbb{R} = \text{card}\,\mathcal{P}(\mathbb{N}).$$

注意, 性质 (b) 的一个重要的特例

$$\text{card}\,(\mathbb{N} \times \mathbb{N}) = \text{card}\,\mathbb{N}$$

可以不用选择公理而只用简单的计数方法直接证明. 由此又容易导出

$$\text{card}\,\mathbb{Q} = \text{card}\,\mathbb{N},$$

也可证明有限个或可数无限个可数无限集的并集是可数无限集.

结合定理 1.5-2 和性质 (c), 可得

$$\text{card}\,\mathbb{N} \prec \text{card}\,\mathbb{R}.$$

所谓连续统假设表述为: 不存在这样的无限集 X, 其基数满足 $\text{card}\,\mathbb{N} \prec \text{card}\,X \prec \text{card}\,\mathbb{R}$. 长期存在的关于连续统假设是否为真的问题在 1963 年和 1964 年被漂亮地解决了, 当时, P. J. Cohen 在两篇里程碑式的文章[5)]里证明了连续统假设独立于 Zermelo-Fraenkel 集合论的六个公理. 而早在 1940 年, K. Gödel 在同样著名的专著[6)] 中证明了: 如果 Zermelo-Fraenkel 集合论并无矛盾, 则添上连续统假设后仍无矛盾.

1.6 拓扑空间

所谓拓扑空间是指一对 $(X, \mathcal{O})$, 其中 X 是一个集合, $\mathcal{O}$ 是 $\mathcal{P}(X)$ 的满足下列条件的一个子集:

对任何一族 $(O_i)_{i\in I}$, 其中 $O_i \in \mathcal{O}$ 其并集 $\cup_{i\in I} O_i \in \mathcal{O}$ (集合 I 可以是有限集也可以是可数或不可数的无限集); 对任何有限个 $(O_j)_{j=1}^n$, 其中 $O_j \in \mathcal{O}$, 其交集 $\cap_{j=1}^n O_j \in \mathcal{O}$; 集合 X 和空集 $\varnothing$ 属于 $\mathcal{O}$.

[5)] P. J. Cohen [1963, 1964]: The independence of the continuum hypothesis. Proceedings of the National Academy of Sciences, USA 50, 1143–1148, and Proceedings of the National Academy of Sciences, USA 51, 105–110.

[6)] K. Gödel [1940]: The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory, Princeton University Press, Princeton, N. J.

如果 $(X,\mathcal{O})$ 是一个拓扑空间, 称集合 X 被赋予一个*拓扑* (对应于 $\mathcal{P}(X)$ 的子集 $\mathcal{O}$), X 中属于 $\mathcal{O}$ 的子集称为*开集* (关于这个拓扑); 设 F 是 X 的子集, 若 $X-F$ 是开集, 则称 F 为*闭集* (关于这个拓扑).

显然, 给定任何一个闭集族 $(F_i)_{i\in I}$, 其交集 $\cap_{i\in I}F_i$ 是闭集; 给定任意有限个闭集 $(F_j)_{j=1}^n$, 其并集 $\cup_{j=1}^n F_j$ 是闭集; 集合 X 和空集是闭集 (这些性质可由 de Morgan 定律导出, 见 1.3 节).

在拓扑空间 $(X,\mathcal{O})$ 中, *点 $x\in X$ 的邻域*是 X 的包含有点 x 的某开子集的任何子集, 点 $x\in X$ 的所有邻域组成的集合记作 $\mathcal{V}(x)$.

给定点 $a,b\in\mathbb{R}$, 满足 $a<b$, 记 $]a,b[:=\{y\in\mathbb{R},\ a<y<b\}$. 一个基本的拓扑空间的例子就是 $(\mathbb{R},\mathcal{O})$, 其中 $\mathbb{R}$ 的非空子集 $O\in\mathcal{O}$ 的充要条件是对任何 $x\in O$, 存在 $a<b$, 使得 $x\in]a,b[$, 而且 $]a,b[\subset O$.

设 $a,b\in\mathbb{R},\ a<b$, 称集合 $]a,b[$ 是以 a,b 为端点的开区间, 它是这个拓扑下的开集. 无界集 $]-\infty,a[:=\{y\in\mathbb{R};y<a\}$, $]b,\infty[:=\{y\in\mathbb{R};b<y\}$ 和 $\mathbb{R}$ 也称作开区间, 它们也是这个拓扑下的开集.

除非有特殊的说明, 集合 $\mathbb{R}$ 上总赋予上述拓扑, 称之为*通常拓扑*. 在这个拓扑中开集的下列特征十分有用.

定理 1.6-1 *设 $\mathbb{R}$ 赋予通常拓扑. 则 $\mathbb{R}$ 中任何非空开子集可表示为有限个或可数无限个互不相交的有界或无界的开区间的并集.*

设 $(X,\mathcal{O})$ 为一个拓扑空间, A 是 X 的一个子集. 称包含于 A 的所有开集的并集为 A 的*内部*, 记作 $\mathring{A}$ 或 $\mathrm{int}\,A$, 即

$$\mathring{A}=\mathrm{int}A:=\{x\in X;\ A\in\mathcal{V}(x)\}.$$

A 的*闭包*是包含 A 的所有闭子集的交集, 记作 $\overline{A}$, 即

$$\begin{aligned}\overline{A}&:=\{x\in X;\ \text{对所有}\ V\in\mathcal{V}(x),\ V\cap A\neq\varnothing\}\\&=X-\{\mathrm{int}\,(X-A)\}.\end{aligned}$$

A 的*边界*定义为 $\overline{A}$ 和 $\overline{X-A}$ 的交集, 记作 ∂A, 即

$$\partial A:=\{x\in X;\ \text{对所有}\ V\in\mathcal{V}(x),\ V\cap A\neq\varnothing\ \text{且}\ V\cap(X-A)\neq\varnothing\}.$$

注意 $\partial A=\overline{A}-\mathring{A}$.

设 $(X,\mathcal{O})$ 为拓扑空间. 实值或复值函数 $f:X\to\mathbb{K}$ 的*支集*是指集合

$$\mathrm{supp}\,f:=\overline{\{x\in X;f(x)\neq 0\}}.$$

设 A 为拓扑空间 $(X,\mathcal{O})$ 的子集, 若 $\overline{A}=X$, 则称 A 在 X 中*稠密*.

拓扑空间 $(X, \mathcal{O})$ 称为可分的, 是指它包含一个有限可数无限的稠密子集, 即存在 $x_n \in X$, $n \geqslant 0$, 使得 $\overline{\cup_{n=0}^{\infty}\{x_n\}} = X$.

设 $(X, \mathcal{O})$ 为拓扑空间, 如果对任意两个不同的点 $x \in X$ 和 $y \in X$, 存在 x 的邻域 V 和 y 的邻域 W, 使得 $V \cap W = \varnothing$, 那么称 X 为 Hausdorff 空间或具有 Hausdorff 拓扑.

拓扑空间 $(X, \mathcal{O})$ 称为正规的, 是指对 X 中任意给定的两个互不相交的闭子集 F_1 和 F_2, 存在互不相交的开子集 O_1 和 O_2, 使得 $F_1 \subset O_1$, $F_2 \subset O_2$.

设 X 为拓扑空间. $x_n \in X$, $n \geqslant 0$. 若存在 $x \in X$, 对 x 的任一邻域 V, 存在整数 $n_0 = n_0(V) \geqslant 0$, 使得当 $n \geqslant n_0$ 时 $x_n \in V$, 则称序列 $(x_n)_{n=0}^{\infty}$ 在 X 中收敛.

如果 X 是 Hausdorff 空间, 上述的 x 便是唯一的. 此时, 记号 $x = \lim\limits_{n\to\infty} x_n$ 或 $x_n \underset{n\to\infty}{\longrightarrow} x$ 或当 $n \to \infty$ 时, $x_n \to x$ 等价地表示为 x 是收敛序列 $(x_n)_{n=0}^{\infty}$ 的极限.

设 X 为一集合, $(Y, \mathcal{O})$ 为 Hausdorff 拓扑空间. 称映射 $f_n : X \to Y$ 的序列 $(f_n)_{n=0}^{\infty}$ 点态收敛于映射 $f : X \to Y$, 是指对每个 $x \in X$, 当 $n \to \infty$ 时, $f_n(x) \to f(x)$.

设 $(X, \mathcal{O})$ 为拓扑空间, A 为 X 的一个子集, 用 $\mathcal{O}_A$ 表示 $\mathcal{P}(A)$ 的一个子集, 其元素为 A 的所有形如 $O_A = \mathcal{O} \cap A$ 的子集, 其中 $O \in \mathcal{O}$. 这样, $(A, \mathcal{O}_A)$ 也是一个拓扑空间, 称这个拓扑是由 $(X, \mathcal{O})$ 的拓扑在 A 上诱导的拓扑, 如果集合 $\mathcal{O}$ 并无歧义, 也简称为诱导拓扑. 由定义可知, A 的属于 $\mathcal{O}_A$ 的子集是诱导拓扑下的开集.

在拓扑空间 $(A, \mathcal{O}_A)$ 中, A 的子集 F_A 为闭集的充要条件是存在 X 的子集 F, 它在拓扑空间 $(X, \mathcal{O})$ 中是闭的, 而且 $F_A = F \cap A$; 类似地, A 的子集 V_A 是点 $x \in A$ 在 $(A, \mathcal{O}_A)$ 中的一个邻域的充要条件是存在 x 在 $(X, \mathcal{O})$ 中的邻域 V, 使得 $V_A = V \cap A$.

当然, 如果 A 是 X 的一个子集, B 是 A 的一个子集, 必须仔细区分 B 在 $(X, \mathcal{O})$ 中的拓扑性质和 B 在 $(A, \mathcal{O}_A)$ 中的拓扑性质. 例如, A 总是 $(A, \mathcal{O}_A)$ 中的开集, 但显然未必是 $(X, \mathcal{O})$ 中的开集.

设 $(X_j, \mathcal{O}_j)$, $1 \leqslant j \leqslant n$, 均为拓扑空间, $X := \prod_{j=1}^{n} X_j$ 表示集合 X_j, $1 \leqslant j \leqslant n$, 的 (有限) 乘积, 记

$$\mathcal{O} := \{O \in \mathcal{P}(X);\ \text{对每个}\ x \in O,\ \text{存在}\ O_j \in \mathcal{O}_j,\ 1 \leqslant j \leqslant n,$$
$$\text{使得}\ x \in O_1 \times \cdots \times O_n,\ \text{且}\ O_1 \times \cdots \times O_n \subset O\}.$$

这样, $(X, \mathcal{O})$ 也是一个拓扑空间, 称 X 上的拓扑是相应于 $\mathcal{P}(x_j)$ 的子集 $\mathcal{O}_j$, $1 \leqslant j \leqslant n$, 的乘积拓扑.

更一般地, 对任意一族拓扑空间 $(X_i, \mathcal{O}_i)$, $i \in I$, 乘积集合 $X = \prod_{i\in I} X_i$ 上的乘积拓扑定义如下: 在此拓扑下子集 $O \subset X$ 称为开集是指对任何 $x \in O$, 存在开集 $O_i \in \mathcal{O}_i$ 的有限族 $(O_i)_{i\in J(x)}$, 使得

$$x \in \left(\prod_{i\in J(x)} O_i\right) \times \left(\prod_{i\in I(x)} X_i\right) \quad \text{且} \quad \left(\prod_{i\in J(x)} O_i\right) \times \left(\prod_{i\in I(x)} X_i\right) \subset O,$$

其中 $I(x) := I - J(x)$.

为使记号简明起见, 对拓扑空间 $(X, \mathcal{O})$, 如果对子集 $\mathcal{O} \subset \mathcal{P}(X)$ 的性质并无歧义, 将不再列出集合 $\mathcal{O}$.

1.7 拓扑空间中的连续性

设 X, Y 均为拓扑空间, $f: X \to Y$ 为从 X 到 Y 的映射, $x \in X$. 若对 Y 中 $f(x)$ 的任何邻域 V, 存在 X 中 x 的一个邻域 U, 使得 U 在 f 下的直接像 $f(U)$ (1.2 节) 包含于 V, 则称 f *在点 x 处连续*.

Hausdorff 空间之间连续映射的一个基本性质是 "它们将收敛序列映射为收敛序列".

定理 1.7-1 *设 X, Y 为两个 Hausdorff 空间, 映射 $f: X \to Y$ 在点 $x \in X$ 处连续, 则对 X 中收敛于点 x 的任何序列 $(x_n)_{n=0}^{\infty}$, 序列 $(f(x_n))_{n=0}^{\infty}$ 在 Y 中收敛于 $f(x)$.*

上述定理的逆命题在 X 为距离空间这个重要的特殊情况下成立 (定理 1.11-1).

可直接证明下列常用的关于复合映射在一点的连续性.

定理 1.7-2 *设 X, Y, Z 为三个拓扑空间, 映射 $f: X \to Y$ 在点 $x \in X$ 处连续, 映射 $g: Y \to Z$ 在点 $f(x) \in Y$ 处连续, 则复合映射 $g \circ f: X \to Z$ 在 x 处连续.*

若映射 $f: X \to Y$ 在 X 上每一点处均连续, 则称 f 是 X 上的连续映射. 可直接证明下列关于连续映射的基本特征:

定理 1.7-3 *设 X, Y 为两个拓扑空间, 映射 $f: X \to Y$ 是连续的充要条件为 Y 中任何开集在 f 下的逆像是 X 中的开集; 或等价于 Y 中任何闭集在 f 下的逆像是 X 中的闭集.*

X 到 Y 的所有连续映射组成的集合记作

$$\mathcal{C}(X; Y) \text{ 或当 } Y = \mathbb{R} \text{ 时, } \mathcal{C}(X).$$

设 X 和 Y 为两个拓扑空间, 称映射 $f: X \to Y$ 是 X 到 Y 上的同胚, 是指 f 为双射, 且 $f \in \mathcal{C}(X; Y)$, $f^{-1} \in \mathcal{C}(Y; X)$.

由定义和定理 1.7-3, 即可得到同胚的下述特征.

定理 1.7-4 *设 X, Y 为两个拓扑空间, $f \in \mathcal{C}(X; Y)$ 是一个双射. 则 f 是 X 到 Y 上的同胚当且仅当 X 中任何开子集在 f 下的直接像是 Y 中的开子集; 或等价地, 当且仅当 X 中任何闭子集在 f 下的直接像是 Y 中的闭子集.*

如果存在拓扑空间 X 到 Y 上的同胚映射, 则称 X 到 Y 是*同胚的*.

读者可以验证当 X 或 Y 为有限乘积时的下述结果.

定理 1.7-5 设 X_j, $1 \leqslant j \leqslant n$, 和 Y 均为拓扑空间, 乘积空间 $X := \prod_{j=1}^n X_j$ 上装备乘积拓扑 (1.6 节), 映射 $f : X \to Y$ 在点 $a = (a_j)_{j=1}^n \in X$ 连续, 则对每个 $1 \leqslant j \leqslant n$, 映射

$$x_j \in X_j \to f\ (a_1, \cdots, a_{j-1}, x_j, a_{j+1}, \cdots, a_n) \in Y$$

在点 a_j 处连续.

注意, 上述定理的逆命题未必成立. 例如, $n = 2, X_1 = X_2 = Y = \mathbb{R}$, 当 $(x_1, x_2) \neq (0,0)$ 时, $f(x_1, x_2) = \dfrac{x_1 x_2}{x_1^2 + x_2^2}$, 当 $(x_1, x_2) = (0,0)$ 时, $f(x_1, x_2) = 0$, $(a_1, a_2) = (0,0)$ 就是一个反例.

定理 1.7-6 设 X 和 Y_i, $1 \leqslant i \leqslant m$, 均为拓扑空间, 乘积空间 $Y := \prod_{i=1}^m Y_i$ 上装备乘积拓扑, 则映射 $f = (f_i)_{i=1}^m : X \to Y$ 在点 $a \in X$ 处连续, 当且仅当每个映射 $f_i : X \to Y_i$, $1 \leqslant i \leqslant m$ 在点 $a \in X$ 处连续.

下述关于连续函数的扩张定理是基本的. 特别地, 它在第 9 章中, 如对定义 $\mathbb{R}^n$ 中 Brouwer 拓扑度, 建立 hairy 球定理和 Borsuk-Ulam 定理等方面有丰富的应用.

定理 1.7-7 (Tietze-Urysohn 扩张定理) 设 X 为正规的拓扑空间, F 为 X 的闭子集, $f : F \to \mathbb{R}$ 为连续函数, 则存在连续函数 $\tilde{f} : X \to \mathbb{R}$, 使得对所有的 $x \in F$,

$$\tilde{f}(x) = f(x).$$

最后, 我们要提及利用一个集合到拓扑空间上给定的映射, 构造这个集合上一类特殊拓扑的基本途径.

定理 1.7-8 设给定集合 X 和从 X 到拓扑空间 Y_i 上映射 φ_i 的族 $(\varphi_i)_{i \in I}$, 则存在 X 上具有下述两个性质的拓扑:

首先, 所有的映射 $\varphi_i : X \to Y_i$, $i \in I$, 均是连续的.

其次, X 上关于这个拓扑的开子集, 关于 X 上使所有映射 $\varphi_i : X \to Y_i$, $i \in I$, 均连续的拓扑也是开的.

由于以上两个性质, 在定理 1.7-8 中定义的拓扑 (显然是唯一的) 称为 X 上使映射 $\varphi_i : X \to Y_i$, $i \in I$, 均连续的最弱拓扑. 我们将会看到赋范向量空间和其对偶空间上的弱拓扑及弱 * 拓扑就是最弱拓扑的基本的例子 (5.12 节).

1.8 拓扑空间中的紧性

设 $(X, \mathcal{O})$ 为拓扑空间, K 为 X 的一个子集, 如果对 X 中任何一族开集 $(O_i)_{i \in I}$, 当 $K \subset \cup_{i \in I} O_i$ 时, 必存在 $(O_i)_{i \in I}$ 的有限子族 $(O_i)_{j \in J}$, 使得 $K \subset \cup_{j \in J} O_j$, 则称 K

为紧集.

上述性质, 即所谓 Heine-Borel-Lebesgue 性质, 常被表述为: 若 X 的子集 K 的任何开覆盖均有有限子覆盖, 则 K 为紧集.

X 的子集 K 为紧集并不取决于把 K 视为 X 的子集或把 K 视为装备诱导拓扑的拓扑空间. 换言之, 拓扑空间 $(X, \mathcal{O})$ 的子集 K 为紧集当且仅当装备诱导拓扑 (1.6 节) 的拓扑空间 $(K, \mathcal{O}_K)$ 是紧的.

下列定理给出了有关紧性的基本性质, 它们都不难验证.

定理 1.8-1 拓扑空间 X 是紧的充要条件是对 X 中闭子集 F_i 组成的任何闭集族 $(F_i)_{i\in I}$, 如果对 $(F_i)_{i\in I}$ 的任何有限子族 $(F_j)_{j\in J}$ 均有 $\bigcap_{j\in J} F_j \neq \varnothing$, 那么必有 $\bigcap_{i\in I} F_i \neq \varnothing$.

定理 1.8-2 (a) Hausdorff 空间的任何紧子集均为闭集.

(b) 紧拓扑空间的闭子集为紧集.

定理 1.8-3 设 X, Y 为两个拓扑空间, $f : X \to Y$ 为连续映射, 则 X 中任何紧集 K 的直接像 $f(K)$ 是 Y 中的紧集.

定理 1.8-4 紧拓扑空间 X 到拓扑空间 Y 上的连续双射是 X 到 Y 上的同胚 (于是, 由定理 1.8-3, Y 也是紧的).

定理 1.8-5 设 X_j, $1 \leqslant j \leqslant n$, 均为紧拓扑空间, 则装备乘积拓扑 (1.6 节) 的乘积空间 $\prod_{j=1}^n X_j$ 是紧空间.

定理 1.8-5 是一般拓扑学中最主要结果之一的 Tychonoff 定理的一个特殊情况. 这一定理断言对任何一族紧的拓扑空间 $(X_i)_{i\in I}$, 装备乘积拓扑的乘积空间 $\prod_{i\in I} X_i$ 是紧的[7]. 但是, 与定理 1.8-5 不同的是后者的证明要用到选择公理.

设 A 为拓扑空间 X 的子集, 如果闭包 $\overline{A}$ 是 X 的紧子集, 则称 A 为相对紧的.

1.9 拓扑空间中的连通和单连通性

拓扑空间 $(X, \mathcal{O})$ 称为连通的, 是指 X 中既是开集又是闭集的子集只有 X 和 $\varnothing$. X 的子集 A 称为连通的, 是指 A 关于 X 在 A 上的诱导拓扑, 它是一个连通的拓扑空间.

X 的子集 A 的连通性并不取决于视 A 为 X 的子集或装备诱导拓扑的拓扑空间.

---

[7] 这个定理首先是对所有的 $i \in I$, $X_i = [0, 1]$ 的特例证明:

A. Tychonoff [1930]: Über die topologische Erweiterung von Räumen. Mathematische Annalen 102, 544–561.

对一般情况的证明见:

E. Čech [1937]: On bicompact spaces. Annals of Mathematics 38, 823–844.

下面的定理罗列了关于连通性的基本性质.

定理 1.9-1 设 A 为拓扑空间 X 的连通子集, 则 X 中满足 $A \subset B \subset \overline{A}$ 的子集 B 也是连通的, 特别地, $B = \overline{A}$ 是连通的.

定理 1.9-2 设 X 为连通的拓扑空间, Y 为拓扑空间, 则任何局部常值函数 $f : X \to Y$ 必为常值函数, 所谓局部常值函数是指对每个点 $x \in X$, 存在邻域 V_x, 使得 $f|_{V_x}$ 为取常值的函数.

类似于紧性, 连通性是"被连续映射所保持"的性质.

定理 1.9-3 设 X, Y 为两个拓扑空间, $f : X \to Y$ 为连续映射, 则 X 中连通子集 A 的直接像 $f(A)$ 是 Y 中的连通子集.

下面的定理给出了 $\mathbb{R}$ 中连通子集的特征.

定理 1.9-4 关于 $\mathbb{R}$ 上的通常拓扑, $\mathbb{R}$ 的子集为连通的当且仅当它是区间 (有界或无界).

定理 1.9-3 和 1.9-4 有以下的直接推论:

定理 1.9-5 (Bolzano 中值定理) 设 X 为连通的拓扑空间, $f : X \to \mathbb{R}$ 为连续函数, $a, b \in X$ 且 $f(a) < f(b)$, 则对任意给定的 $y \in]f(a), f(b)[$, 存在 $x \in X$ 使得 $f(x) = y$.

以下三个定理给出了关于连通性的常用的充分条件.

定理 1.9-6 设 X 为拓扑空间, $(A_i)_{i \in I}$ 为 X 的一族连通子集. 若交集 $\bigcap_{i \in I} A_i$ 非空, 则其并集 $\cup_{i \in I} A_i$ 是连通的.

定理 1.9-7 设 X_j, $1 \leqslant j \leqslant n$, 均为连通的拓扑空间, 则其装备乘积拓扑 (1.6 节) 的乘积空间 $X = \prod_{j=1}^{n} X_j$ 是连通的.

设 X 为拓扑空间. 关系 $\mathcal{R}$ 定义为"$(x, y) \in \mathcal{R}$ 当且仅当存在 X 中包含 x 与 y 的连通子集"是 X 中的一个等价关系. 模这个关系的等价类为 X 的子集, 称为 X 的*连通分支*.

给定 $x \in X$, X 中包含 x 的连通分支称为 *x 的连通分支*; 根据下述结果, 它是 X 中包含 x 的最大连通子集.

定理 1.9-8 设 X 为拓扑空间, $x \in X$. 则 x 的连通分支即 X 包含 x 的所有连通子集的并.

设 x 和 y 是拓扑空间 X 中两个点. 所谓*联结 x 到 y 的道路*是指连续映射 $\gamma : [0, 1] \to X$, 它满足 $\gamma(0) = x$, $\gamma(1) = y$.

如果对拓扑空间 X 中任何两个不同的点, 均存在联结 x 到 y 的道路, 那么称 X

为弧连通 (或道路连通) 的.

定理 1.9-9 弧连通的拓扑空间是连通的.

这个结论的逆并不成立. 例如, 设

$$A := \{(x,y) \in \mathbb{R}^2; x = 0, |y| \leqslant 1\}, B := \left\{(x,y) \in \mathbb{R}^2; x \neq 0,\ y = \sin\frac{1}{x}\right\},$$

则 $A \cup B$ 是 $\mathbb{R}^2$ 的连通子集, 但它却不是弧连通的.

设 x 和 y 是拓扑空间 X 中的两个点, 联结 x 到 y 的两条道路 $\gamma_0 : [0,1] \to X$ 和 $\gamma_1 : [0,1] \to X$ 称为同伦的, 是指存在连续映射 $H : [0,1] \times [0,1] \to X$, 满足 $H(\cdot, 0) = \gamma_0$, $H(\cdot, 1) = \gamma_1$, 且 $H(0,\cdot) = x$, $H(1,\cdot) = y$, 称 H 为联结 γ_0 到 γ_1 的同伦.

拓扑空间 X 称为单连通的, 是指它是弧连通的, 因而是连通的 (定理 1.9-9), 而且对任何两条道路 $\gamma_0 : [0,1] \to X$ 和 $\gamma_1 : [0,1] \to X$, 当 $\gamma_0(0) = \gamma_1(0)$, $\gamma_0(1) = \gamma_1(1)$ 时, 它们必是同伦的.

1.10 距离空间

设 X 为一个集合, 函数 $d : X \times X \to \mathbb{R}$ 满足下列性质, 对任何 $x, y, z \in X$:

$$d(x,x) = 0, \text{ 且当 } x \neq y \text{ 时, } d(x,y) > 0,$$
$$d(x,y) = d(y,x),$$
$$d(x,z) \leqslant d(x,y) + d(y,z),$$

则称 d 为 X 上的距离. 上述最后一条性质称为三角不等式. 设 X 为一集合, d 为 X 上的距离, 称 (X,d) 为距离空间.

以下总用 (X,d) 表示距离空间. 给定 $x \in X$ 和数 $r > 0$, 定义 X 的子集

$$B(x;r) = \{y \in X; d(y,x) < r\}$$

为以 x 为中心或中心在 x 处, 半径为 r 的球.

设 A 为 X 的子集, 如果存在一个球 $B(x;r) \subset X$, 使得 $A \subset B(x;r)$, 则称 A 是有界的. 否则, 称 A 为无界的.

X 的非空子集 A 的直径定义为一个扩张的实数

$$\operatorname{diam} A := \sup\{d(x,y);\ x \in A, y \in A\} \in [0, +\infty].$$

点 $x \in X$ 到 X 的非空子集 A 的距离定义为实数

$$\operatorname{dist}(x, A) := \inf\{d(x,y); y \in A\}.$$

除非特别说明, 距离空间 (X,d) 总可以视为拓扑空间 $(X,\mathcal{O})$, 其中的开集, 即 X 中属于 $\mathcal{O}$ 的子集由下述定理给出. 这个典则拓扑称为由距离 d 在 X 上诱导的拓扑.

定理 1.10-1 设 (X,d) 为距离空间, 用 $\mathcal{O}$ 表示由空集和 X 的所有满足下述性质的子集 O 组成的 $\mathcal{P}(X)$ 的子集: 给定任何 $x\in O$, 存在 $r>0$, 使得球 $B(x;r)$ 包含于 O. 则 $(X,\mathcal{O})$ 是 Hausdorff 的而且是正规的拓扑空间.

对 $x\in X$, $r>0$, 球 $B(x;r)$ 在此拓扑下为开集.

例如, 在 $\mathbb{R}$ 上由对任何 $x,y\in\mathbb{R}$, 令 $d(x,y)=|x-y|$ 定义的通常距离, 导出 $\mathbb{R}$ 上的通常拓扑 (1.6 节). 类似地, 在 $\mathbb{C}$ 上由对任何 $x,y\in\mathbb{C}$, 令 $d(x,y)=|x-y|$ 定义的通常距离在 $\mathbb{C}$ 上导出的拓扑称为 $\mathbb{C}$ 上的通常拓扑.

将集合 $\mathbb{R}$ 或 $\mathbb{R}$ 的一个子集视作距离空间时, 总隐含着取 $\mathbb{R}$ 上的通常距离 d.

当然, d 远非定义通常拓扑的唯一距离. 例如, 对任何 $x,y\in\mathbb{R}$, 由 $\rho(x,y)=\dfrac{|x-y|}{1+|x-y|}$ 定义了距离 $\rho:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$, 它也在 $\mathbb{R}$ 上导出同样的通常拓扑. 然而, 要注意距离空间 $(\mathbb{R},d)$ 是无界的, 而距离空间 $(\mathbb{R},\rho)$ 是有界的. 这个简单的例子也说明有界性是一个距离的概念, 而并非拓扑概念.

如果拓扑空间 $(X,\mathcal{O})$ 的拓扑可由一个 X 上的距离导出, 则称此拓扑是可距离化的, 称这个距离与此拓扑相容. 例如, $\mathbb{R}$ 上的通常拓扑是可距离化的, 前面提及的距离 d 和 ρ 均与之相容.

距离空间另一个基本的例子是 $\mathbb{K}^n$ 或 $\mathbb{K}^n$ 的子集 X, 这里 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$, 其上装备的距离 d_p, $1\leqslant p\leqslant+\infty$ 定义为对任何 n 维元 $x=(x_i)_{i=1}^n\in\mathbb{K}^n$ 和 $y=(y_i)_{i=1}^n\in\mathbb{K}^n$,

$$d_p(x,y)=\left(\sum_{i=1}^n|x_i-y_i|^p\right)^{\frac{1}{p}}, \text{ 其中 } 1\leqslant p<+\infty,$$

$$d_\infty(x,y)=\max_{1\leqslant i\leqslant n}|x_i-y_i|,$$

除了对当 $1<p<+\infty$ 时的三角不等式而外, 容易验证上述 d_p 满足距离的所有公理, 前者可参见定理 2.4-1. 距离 d_2 称为 Euclid 距离.

给定任何 $1\leqslant p,\ q\leqslant\infty$, 对应于距离 d_p 的任何一个球均可包含于一个中心相同的对应于距离 d_q 的球. 因此, 由任何一个 d_p, $1\leqslant p\leqslant+\infty$ 在 $\mathbb{K}^n$ 或 $\mathbb{K}^n$ 的子集上导出的拓扑均是相同的, 则称之为 $\mathbb{K}^n$ 上的通常拓扑. 容易验证这个拓扑与 $\mathbb{K}^n=\prod_{j=1}^n X_j$ 上的乘积拓扑 (1.6 节) 也是一致的, 这里每个拓扑空间 X_j, $1\leqslant j\leqslant n$ 即装备通常拓扑的集合 $\mathbb{K}$. 关于这个拓扑, $\mathbb{K}^n$ 是可分拓扑空间.

更一般地, 对于 $\mathbb{K}$ 上以 $(e_i)_{i=1}^n$ 为基的有限维向量空间的子集 X, 将上述定义 $d_p(x,y)$ 中的 x,y 用 $x=\sum_{i=1}^n x_ie_i$, $y=\sum_{i=1}^n y_ie_i$ 取代后, 按 d_p, $1\leqslant p\leqslant+\infty$ 成为一个距离空间.

为记号简单起见, 以下将距离空间 (X,d) 简记为 X, 这里视需要隐含着记距离

为 d.

在 1.6 节中给出的一些定义, 在距离空间 X 中有下列等价的用距离的表述:

点 $x \in X$ 的邻域是 X 的任何包含一个以 x 为中心的球的子集.

X 的子集 A 的内部 $\overset{\circ}{A}$ 是存在以 x 为中心且包含于 A 的球的点 $x \in A$ 的集合.

X 的子集 A 的闭包 $\overline{A}$ 是任何以 x 为中心的球与 A 均有非空交的点 $x \in X$ 的集合.

X 的子集 A 的边界 ∂A 是任何以 x 为中心的球与 A 及 $X - A$ 均有非空交的点 $x \in X$ 的集合.

距离空间 (X, d) 称为可分的, 是指存在 $x_n \in X, n \geqslant 1$, 使得对任何 $x \in X$ 和任意的 $\varepsilon > 0$, 存在 $n = n(x, \varepsilon) \geqslant 1$ 满足 $x \in B(x_n; \varepsilon)$.

点 $x_n \in X$ 组成的序列 $(x_n)_{n=0}^{\infty}$ 收敛是指存在点 x 使得对任何 $\varepsilon > 0$, 存在 $n_0 \geqslant 0$, 当 $n \geqslant n_0$ 时, $x_n \in B(x; \varepsilon)$; 或等价于当 $n \to \infty$ 时, $d(x_n, x) \to 0$. 由于相应的拓扑空间是 Hausdorff 空间 (定理 1.10-1), 所以上述的 x 是唯一的, 即这个序列的极限.

利用极限可用距离描述, 距离空间中子集的闭包也可用收敛序列给出简单的特征.

定理 1.10-2 设 A 为距离空间的一个子集, 则点 $x \in X$ 属于 $\overline{A}$ 当且仅当存在 $x_n \in A$ 的序列 $(x_n)_{n=0}^{\infty}$, 当 $n \to \infty$ 时, 收敛于 x.

由此可知, X 的子集 A 为闭的, 当且仅当

$$x_n \in A,\ n \geqslant 0 \text{ 且当 } n \to \infty \text{ 时 } x_n \to x \in X, \text{ 则 } x \in A.$$

顺便注意到如果 A 是闭集, $x \overline{\in} A$, 则 $\operatorname{dist}(x, A) > 0$.

设 (X, d) 为距离空间, A 为 X 的一个子集, 用 $d_A : A \times A \to \mathbb{R}$ 表示距离 d 在 $A \times A$ 上的限制. 显然, d_A 是 A 上的一个距离, 称为由 d 在 A 上导出的距离, 这样 (A, d_A) 也是一个距离空间, 而且, 有下述性质.

定理 1.10-3 设 (X, d) 为距离空间, A 为 X 的一个子集. 由距离 d_A 在 A 上导出的拓扑与 X 上由 d 导出的拓扑在 A 上导出的拓扑 (1.6 节) 相一致.

并且, 若 (X, d) 可分, 则 (A, d_A) 可分.

最后, 设 (X_j, d_j), $1 \leqslant j \leqslant n$, 均为距离空间, $X := \prod_{j=1}^{n} X_j$. 则在乘积空间 X 的乘积拓扑下, 子集 $O \subset X$ 为开集的充要条件是对任何 $x = (x_j)_{j=1}^{n} \in X$, 存在 $r_j > 0$, $1 \leqslant j \leqslant n$, 使得 $\prod_{j=1}^{n} B(x_j; r_j) \subset O$. 在乘积空间 X 上任何导出乘积拓扑的距离均称为与乘积拓扑相容. 如函数 $d : X \times X \to \mathbb{R}$ 和 $\rho : X \times X \to \mathbb{R}$ 分别定义为

$$d(x, y) := \sum_{j=1}^{n} d_j(x_j, y_j) \quad \text{和} \quad \rho(x, y) := \max_{1 \leqslant j \leqslant n} d_j(x_j, y_j),$$

其中 $x = (x_j)_{j=1}^{n} \in X$, $y = (y_j)_{j=1}^{n} \in X$, 就是相容距离的例子.

1.11 距离空间的连续性和一致连续性

从拓扑空间 X 到拓扑空间 Y 的映射若在点 $x \in X$ 处连续, 则它将在 X 中收敛于点 x 的序列映射为 Y 中收敛于点 $f(x)$ 的序列 (定理 1.7-1). 当 X 与 Y 的拓扑由距离导出时这个命题的逆也是成立的 (实际上, 当 X 为距离空间, Y 为 Hausdorff 拓扑空间时, 定理 1.11-1 依然成立):

定理 1.11-1 设 X, Y 均为距离空间, 则映射 $f: X \to Y$ 在点 $x \in X$ 处连续, 当且仅当对 X 中任何一列收敛于 x 的序列 $(x_n)_{n=0}^{\infty}$, Y 中的序列 $(f(x_n))_{n=0}^{\infty}$ 均收敛于 $f(x)$.

下面的定理给出了距离空间中连续映射的一个简单而常用的性质.

定理 1.11-2 设 X 为距离空间 $\widetilde{X}$ 的一个稠密子集, Y 为 Hausdorff 拓扑空间, $f: \widetilde{X} \to Y$ 与 $g: \widetilde{X} \to Y$ 是两个连续映射, 且在 X 上相同, 即对所有 $x \in X$, 有 $f(x) = g(x)$, 则 $f = g$.

如果 X 和 Y 均为距离空间, 则映射在某点处的连续性可以用球或距离来描述: 设 (X, d) 和 (Y, ρ) 为距离空间, 则映射 $f: X \to Y$ 在点 $x \in X$ 处连续, 是指 Y 中任何以 $f(x)$ 为中心的球的原像包含一个 X 中以 x 为中心的球, 这又等价于对任意 $\varepsilon > 0$, 存在 $\delta = \delta(\varepsilon, x) > 0$, 使得 $\tilde{x} \in X$, 当 $d(x, \tilde{x}) < \delta$ 时 $\rho(f(x), f(\tilde{x})) < \varepsilon$. 距离空间上这一等价定义有时也称作 "连续性的 $\varepsilon - \delta$ 定义".

如果映射 $f: X \to Y$ 是连续的, 即在任何 $x \in X$ 处均连续, 可能会发生这样的情况: 对任意 $\varepsilon > 0$, 上述的数 $\delta(\varepsilon, x) > 0$ 可与 $x \in X$ 无关地选取到. 这种可能性导致下述的定义: 设 (X, d) 和 (Y, ρ) 为两个距离空间, 映射 $f: X \to Y$. 如果对任意 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$, 当 $x, \tilde{x} \in X$ 时, $d(x, \tilde{x}) < \delta(\varepsilon)$, 那么称 f 是一致连续的.

一致连续映射的一个重要例子是所谓 Lipschitz 连续映射, 即对于映射 $f: (X, d) \to (Y, \rho)$, 存在常数 k, 使得

$$\rho(f(x), f(\tilde{x})) \leqslant k d(x, \tilde{x})$$

对一切 $x, \tilde{x} \in X$ 成立. 这样的映射也称作满足带 Lipschitz 常数 k 的 Lipschitz 条件.

另一个例子是 Hölder 连续映射, 即映射 $f: (X, d) \to (Y, \rho)$ 具有性质: 存在常数 C 和 $0 < \lambda < 1$, 使得对任何 $x, \tilde{x} \in X$,

$$\rho(f(x), f(\tilde{x})) \leqslant C(d(x, \tilde{x}))^{\lambda}$$

成立, 这样的映射称为满足指数为 λ 的 Hölder 条件.

设 (X, d) 为距离空间, 在乘积空间 $X \times X$ 上定义距离 D 为对任何 $(x, \tilde{x}) \in X \times X$ 和 $(y, \tilde{y}) \in X \times X$,

$$D((x, \tilde{x}), (y, \tilde{y})) := d(x, y) + d(\tilde{x}, \tilde{y}).$$

这样, 函数 $d:(X\times X,D)\to\mathbb{R}$ 就是带有 Lipschitz 常数 1 的 Lipschitz 连续函数. 为说明这一点, 只要注意到三角不等式

$$|d(x,\tilde{x})-d(y,\tilde{y})|\leqslant d(x,y)+d(\tilde{x},\tilde{y})=D((x,\tilde{x}),(y,\tilde{y}))$$

即可.

另一个类似的例子由到子集的距离给出:

定理 1.11-3 设 (X,d) 为距离空间, A 为 X 的非空子集, 则对任何 $x,y\in X$,

$$|\text{dist }(x,A)-\text{dist }(y,A)|\leqslant d(x,y).$$

给定两个距离空间 (X,d) 和 (Y,ρ), 如果 X 到 Y 的映射 $f:X\to Y$ "保持距离", 即对任何 $x,y\in X$, 均有 $\rho(f(x),f(y))=d(x,y)$, 则称 f 为 X 到 Y 的等距映射. 等距映射给出了一致连续映射的另一个例子.

此外, 一个较一般而常用的一致连续性的充分条件将在定理 1.13-2 中给出.

1.12 完备距离空间

设 (X,d) 为距离空间, $x_n\in X$, $n\geqslant 0$. 如果当 $n\to\infty$ 时, 集合 $\cup_{m=n}^{\infty}\{x_m\}$ 的直径 (1.10 节) 收敛于 0, 或等价地, 对任何 $\varepsilon>0$, 存在 $n_0(\varepsilon)\geqslant 0$, 使得当 $m\geqslant n_0(\varepsilon)$, $n\geqslant n_0(\varepsilon)$ 时, $d(x_m,x_n)<\varepsilon$, 就称序列 $(x_n)_{n=0}^{\infty}$ 为 Cauchy 序列.

下面的定理列出了 Cauchy 序列的基本性质:

定理 1.12-1 (a) Cauchy 序列是有界的.

(b) 收敛序列是 Cauchy 序列.

(c) 包含一收敛子列的 Cauchy 序列是收敛的, 其极限即子列的极限.

距离空间 (X,d) 为完备的是指 X 中的每个 Cauchy 序列均在 X 中收敛. 距离空间的子集 A 为完备的是指距离空间 (X,d_A) 是完备的, 其中 d_A 为 d 在 A 上导出的距离 (1.10 节). 由此, "X 为完备距离空间" 这一性质与 X 是否为一个更大的距离空间的子集无关.

下面的定理列出了完备距离空间的基本性质:

定理 1.12-2 设 A 为距离空间 X 的子集.

(a) 若 A 也是完备的, 则 A 在 X 中是闭的.

(b) 若 X 是完备的, A 在 X 中是闭的, 则 A 是完备的.

(c) 若 X 是完备的, X 的子集 A 是完备的当且仅当 A 在 X 中是闭的.

完备距离空间的例子如装备通常距离的 $\mathbb{R}$ 和 $\mathbb{C}$; 装备任何一个距离 d_p, $1\leqslant p\leqslant\infty$ (1.10 节) 的 $\mathbb{R}^n$ 和 $\mathbb{C}^n$, $n\geqslant 2$: $\mathbb{R}$ 和 $\mathbb{C}$ 的完备性由其构造 (1.4 节) 可得; $(\mathbb{R},d_p)$ 和 $(\mathbb{C}^n,d_p)$ 的完备性容易从 $\mathbb{R}$ 和 $\mathbb{C}$ 的完备性导出.

对给定的 $n \geqslant 2$, 距离 d_p, $1 \leqslant p \leqslant +\infty$ 给出了在 $\mathbb{R}^n$ 上导出相同的拓扑, 同时 $(\mathbb{R}^n, d_p)$ 又均是完备的距离的例子. 然而一般说来并非如此. 例如, 设 $\mathbb{R}_+ := \{x \in \mathbb{R}; x \geqslant 0\}$, 在 $\mathbb{R}_+$ 上的距离 d 和 ρ 分别定义为 $d(x,y) = |x-y|$, $\rho(x,y) = |\frac{x}{1+x} - \frac{y}{1+y}|$, 其中 $x, y \in \mathbb{R}_+$. 这两个距离在 $\mathbb{R}_+$ 上诱导出相同的拓扑, 但 $(\mathbb{R}_+, d)$ 是完备的, 而 $(\mathbb{R}_+, \rho)$ 却不完备 (取 $x_n = n$, 则 $(x_n)_{n=0}^\infty$ 是 $(\mathbb{R}_+, \rho)$ 中的 Cauchy 序列, 但在 $(\mathbb{R}_+, \rho)$ 中并不收敛).

下面的定理是基本的. 它提供了一个定义于某距离空间稠密子空间上的连续映射可以延拓为全空间上连续映射的充分条件 (这一结果将在定理 3.1-1 中对赋范向量空间加以证明).

定理 1.12-3 (唯一连续延拓) 设 X 为距离空间 $\widetilde{X}$ 的稠密子空间, Y 为完备距离空间, $f: X \to Y$ 为一致连续映射.

则映射 f 存在唯一的到全空间 $\widetilde{X}$ 的连续延拓 $\tilde{f}: \widetilde{X} \to Y$. 映射 $\tilde{f}$ 在 $\widetilde{X}$ 上也是一致连续的.

下面的定理也是基本的. 它指出任何不完备的距离空间可以在等距的意义下等同于一个完备距离空间的稠密子集 (这个结果也将在定理 3.1-2 中对赋范向量空间加以证明).

定理 1.12-4 (距离空间完备化) (a) 设 (X, d) 为距离空间, 则存在完备距离空间 $(\widetilde{X}, \tilde{d})$ 和等距映射 $\sigma: X \to \widetilde{X}$, 使 $\sigma(X)$ 在 $\widetilde{X}$ 中稠密.

(b) 若空间 X 是可分的, 则空间 $\widetilde{X}$ 也是可分的.

(c) 若 $(\widehat{X}, \hat{d})$ 是完备距离空间, 且存在 X 到 $\widehat{X}$ 的一个稠密子集上的等距映射, 则必存在一个从 $(\widetilde{X}, \tilde{d})$ 到 $(\widehat{X}, \hat{d})$ 上的等距映射.

由于性质 (c), 作为距离空间, $(\widetilde{X}, \tilde{d})$ 在等距双射的意义下是唯一的. 这个 "本质上唯一" 的空间称为距离空间 (X, d) 的完备化空间.

另外两个关于完备距离空间的基本定理, 即 Banach 不动点定理和 Baire 定理将在后面章节加以证明 (定理 3.7-1 和定理 5.1-2).

1.13 距离空间中的紧性

设 (X, d) 为距离空间, K 为 X 的子集, K 上的拓扑由距离 d 导出. 如果 K 作为拓扑空间, 满足 Heine-Borel-Lebesgue 性质 (1.8 节), 则 K 是紧的.

注意到 Hausdorff 拓扑空间的任何紧子集是闭的 (定理 1.8-2(a)); 在紧距离空间中任何半径为 1 的球的任何覆盖必有有限子集覆盖 (由 Heine-Borel-Lebesgue 性质可得), 于是即可导出紧性的一个必要条件:

定理 1.13-1 距离空间的紧子集必是闭的, 且是有界的.

距离空间中 Heine-Borel-Lebsegue 性质的另一个简单的推论是关于一致连续性的一个充分条件.

定理 1.13-2 设 X 为紧距离空间, Y 为距离空间, 则 X 到 Y 的连续映射必是一致连续的.

设 A 为距离空间 X 的一个子集. 如果对任何 $\varepsilon > 0$, 存在有限个点 $x_j = x_j(\varepsilon) \in A$, $1 \leqslant j \leqslant n$, $n = n(\varepsilon)$, 使得

$$A \subset \bigcup_{j=1}^{n} B(x_j; \varepsilon),$$

则称 A 为准紧的. 注意, $A \subset X$ 是准紧的, 当且仅当 $\overline{A}$ 是准紧的.

以下关于距离空间中紧和准紧子集的特征是基本的.

定理 1.13-3 设 X 为距离空间, K 为 X 的一个子集. 下列三个命题等价:

(a) K 是 X 的紧子集.

(b) 由 K 中的点 x_n 组成的任何序列 $(x_n)_{n=0}^{\infty}$, 均存在收敛于 K 中点的子列 $(x_{\sigma(n)})_{n=0}^{\infty}$.

(c) K 是准紧的且是完备的.

满足上述性质 (b) 的拓扑空间 A 称为满足 Bolzano-Weierstrass 性质, 它推广了定理 1.4-1(b).

定理 1.13-4 距离空间中的子集 A 是相对紧集的充要条件是任何点 $x_n \in A$ 的序列 $(x_n)_{n=0}^{\infty}$ 均包含在 $\overline{A}$ 中收敛的子列 $(x_{\sigma(n)})_{n=0}^{\infty}$.

虽然定理 1.13-1 的逆一般说来并不成立, 但是定理 1.13-3(c) 说明在下列特殊情况下它是成立的.

定理 1.13-5 在空间 $\mathbb{K}^n$ 上装备任何一个距离 d_p, $1 \leqslant p \leqslant \infty$ (1.10 节), 其中 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$, 则 $\mathbb{K}^n$ 的子集为紧集的充要条件是它为有界闭集.

后面我们将证明这个性质实际上是有限维空间的特征. 这就是重要的 F. Riesz 定理 (定理 2.7-3) 的本质所在.

定理 1.13-5 给出了有限维空间紧集的特征, 无限维空间中紧集的特征则是棘手的问题, 一个重要的例子[8] 是紧集上连续函数空间中相应的特征为 Ascoli-Arzelà 定理 (定理 3.10-1).

由定理 1.13-5, $\mathbb{R}$ 的子集为紧集的充要条件是它为有界闭集. 结合定理 1.8-3 可以得到另外一个基本结果, 即紧集上的连续函数可以取到其下确界和上确界:

---

[8] 另一个重要的例子是空间 $L^{p(\Omega)}$ 上的 Kolmogorov 定理, 其证明可见 Brezis [2011] 中的定理 4.26.

定理 1.13-6 设 K 为紧拓扑空间, $f: K \to \mathbb{R}$ 为连续函数, 则 K 的直接像 $f(K)$ 是 $\mathbb{R}$ 中的紧子集, 从而存在 $x_0 \in K$ 和 $x_1 \in K$, 使得

$$f(x_0) = \inf_{x \in K} f(x), \quad f(x_1) = \sup_{x \in K} f(x).$$

下一章将给出定理 1.13-6 的一个奇妙的应用, 即以此给出代数学基本定理 (定理 2.8-1) 的一个简单的证明.

1.14 $\mathbb{R}^n$ 中的 Lebesgue 测度; 可测函数

在本节和下一节中, 用记号 $[0,\infty]$ 和 $[-\infty,\infty]$ 分别表示集合 $[0,\infty[\cup\{\infty\}$ 和 $\{-\infty\}\cup\mathbb{R}\cup\{\infty\}$ (1.4 节).

设 X 为一个集合, $\mathcal{A}$ 为 $\mathcal{P}(X)$ 的一个子集, 满足下列条件:

$$X \in \mathcal{A},$$
$$A \in \mathcal{A} \text{ 蕴含 } (X - A) \in \mathcal{A},$$
$$\text{对 } i \geqslant 1, \text{ 当 } A_i \in \mathcal{A} \text{ 时}, \bigcup_{i=1}^{\infty} A_i \in \mathcal{A},$$

则称 $\mathcal{A}$ 为 X 的子集的一个 σ 代数.

给定集合 X 和 X 的子集的 σ 代数 $\mathcal{A}$, 函数 $\mu: \mathcal{A} \to [0,\infty]$ 满足下列条件:

$$\mu(\varnothing) = 0,$$

如果对 $i \geqslant 1$, $A_i \in \mathcal{A}$, 且 $i \neq j$ 时, $A_j \cap A_j = \varnothing$, 则

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i),$$

那么称 μ 是测度.

以上后一条件也称作测度 μ 的 σ 可加性. 称三元组合 $(X, \mathcal{A}, \mu)$ 为测度空间.

十分基本而重要的是集合 $X = \mathbb{R}^n$, 其中 n 为正整数, 其上赋以通常拓扑 (1.10 节), 相应的 σ 代数取 $\mathbb{R}^n$ 中的 Borel σ 代数 $\widetilde{\mathcal{A}}$, 其定义为包含 $\mathbb{R}^n$ 中所有开子集的最小 σ 代数 (σ 代数 $\widetilde{\mathcal{A}}$, 作为 $\mathbb{R}^n$ 中包含所有开子集的一切 σ 代数的交集, 它是唯一确定的), 相应的测度为 Lebesgue 测度 $\widetilde{\mu}$, 其定义为对任何 $\widetilde{A} \in \widetilde{\mathcal{A}}$,

$$\widetilde{\mu}(\widetilde{A}) = \inf\left\{\sum_{k=1}^{\infty}\left(\prod_{j=1}^{n}(b_j^k - a_j^k)\right); \widetilde{A} \subset \bigcup_{k=1}^{\infty}\left(\prod_{j=1}^{n}\,]a_j^k, b_j^k[\right)\right\}.$$

对于给定的 $\widetilde{A} \in \widetilde{\mathcal{A}}$, 这里的下确界是对所有覆盖 $\widetilde{A}$ 的开区间乘积 $\prod_{j=1}^{n}\,]a_j^k, b_j^k[$, $k \geqslant 1$, 的可数无限族的并来取. 这个 σ 代数的元称为 $\mathbb{R}^n$ 中的 Borel 可测子集.

用这个方法构造的测度空间 $(\mathbb{R}^n, \widetilde{\mathcal{A}}, \widetilde{\mu})$ 缺少一个人们常要求的性质, 即当 $\widetilde{A} \in \widetilde{\mathcal{A}}$ 满足 $\widetilde{\mu}(\widetilde{A}) = 0$ 时, $\widetilde{A}$ 的子集也在 σ 代数 $\widetilde{\mathcal{A}}$ 中. 为了排除这个困难, 令

$\mathcal{A} := \{A \in \mathcal{P}(\mathbb{R}^n)$; 存在 $\widetilde{A} \in \widetilde{\mathcal{A}}$, $\widetilde{A}' \in \widetilde{\mathcal{A}}$, $\widetilde{B} \subset \widetilde{A}'$, 其中 $\widetilde{\mu}(\widetilde{A}') = 0$, 使得 $A = \widetilde{A} \cup \widetilde{B}\}$,

对任何 $A \in \mathcal{A}$, 令

$$\mu(A) := \widetilde{\mu}(\widetilde{A}),$$

其中 $\widetilde{A} \in \widetilde{\mathcal{A}}$ 满足存在 $\widetilde{A}'$ 和 $\widetilde{B}$, 使得 $\mu(\widetilde{A}') = 0$, $\widetilde{B} \subset \widetilde{A}'$, 而且 $A = \widetilde{A} \cup \widetilde{B}$.

可以证明 $\mathcal{A}$ 也是 $\mathbb{R}^n$ 的子集的一个 σ 代数 (它显然包含 Borel σ 代数 $\widetilde{\mathcal{A}}$), 上面关于 $\mu(A)$ 的定义是一意确定的 (即与特殊的 $\widetilde{A} \in \widetilde{\mathcal{A}}$ 的选取无关), 如上定义的函数 $\mu : \mathcal{A} \to [0, \infty]$ 也是一个测度.

称 σ 代数 $\mathcal{A}$ 为 $\mathbb{R}^n$ 上的 Lebesgue σ 代数, 称 $\mathcal{A}$ 的元为 $\mathbb{R}^n$ 中的 Lebesgue 可测子集, 也称 μ 为 $\mathbb{R}^n$ 上的 Lebesgue 测度, 或 n 维 Lebesgue 测度. 显然, 当 $\widetilde{A} \in \widetilde{\mathcal{A}}$ 时, $\mu(\widetilde{A}) = \widetilde{\mu}(\widetilde{A})$.

$\mathbb{R}^n$ 上的 Lebesgue 测度根据上下文记作

$$dx \text{ 或 meas 或 } dx\text{-meas}.$$

关于基数, 可以证明

$$\operatorname{card} \widetilde{\mathcal{A}} = \operatorname{card} \mathbb{R}, \quad \operatorname{card} \mathcal{A} = \operatorname{card} \mathcal{P}(\mathbb{R}).$$

因为 $\operatorname{card} \mathbb{R} \prec \operatorname{card} \mathcal{P}(\mathbb{R})$ (定理 1.5-2), 所以 $\mathbb{R}^n$ 中的 Lebesgue 可测子集比 $\mathbb{R}^n$ 中的 Borel 可测子集 “多得多”.

下面的定理概括了上述的测度空间 $(\mathbb{R}^n, \mathcal{A}, \mu)$ 的四条基本性质. 前三个性质是其构造的直接推论, 第四个性质说明 Lebesgue 测度是平移不变的.

定理 1.14-1　$\mathbb{R}^n$ 中 Lebesgue 可测子集的 σ 代数 $\mathcal{A}$ 和 Lebesgue 测度 $\mu : \mathcal{A} \to [0, \infty]$ 满足下列性质:

(a) $\mathbb{R}^n$ 中的每个开集属于 $\mathcal{A}$; 因此 $\mathbb{R}^n$ 中的每个闭集, 以及 $\mathbb{R}^n$ 中可数个开集或闭集的交集与并集均属于 $\mathcal{A}$.

(b) $\mathbb{R}^n$ 中任何形如 $\prod_{j=1}^n]a_j, b_j[$ 的子集 (由 (a), 它属于 $\mathcal{A}$) 的 Lebesgue 测度为

$$\mu\left(\prod_{j=1}^n]a_j, b_j[\right) = \prod_{j=1}^n (b_j - a_j).$$

(c) 如果 $A \in \mathcal{A}$ 且 $\mu(A) = 0$, 则 A 的任何子集也是 Lebesgue 可测的, 其 Lebesgue 测度为 0.

(d) 对任意给定的 $x \in \mathbb{R}^n$ 和 $A \in \mathcal{A}$, 集合

$$x + A = \{(x + y) \in \mathbb{R}^n;\ y \in A\}$$

也属于 $\mathcal{A}$, 且 $\mu(x+A)=\mu(A)$.

值得注意的是由 Lebesgue 测度的平移不变性 (d), 并利用选择公理可以导出 $\mathbb{R}^n$ 中 Lebesgue 不可测集的存在性.

下面, 我们来说明测度空间 $(X,\mathcal{A},\mu)$ 和 $(Y,\mathcal{B},\nu)$ 的乘积空间也可构作一个测度空间. 首先, 用 $\mathcal{A}\otimes\mathcal{B}$ 表示包含所有集合 $A\times B\in\mathcal{P}(X\times Y)$, 其中 $A\in\mathcal{A}$, $B\in\mathcal{B}$ 的最小 σ 代数 (由这些条件, $\mathcal{A}\otimes\mathcal{B}$ 是唯一确定的). 可以证明存在唯一的乘积测度

$$\mu\otimes\nu:\mathcal{A}\otimes\mathcal{B}\to[0,\infty]$$

满足对任何 $A\in\mathcal{A}$ 和 $B\in\mathcal{B}$, 有

$$(\mu\otimes\nu)(A\times B)=\mu(A)\nu(B).$$

可以证明, 当 μ 是 $\mathbb{R}^m$ 上的 Lebesgue 测度, ν 是 $\mathbb{R}^n$ 上的 Lebesgue 测度时, 乘积测度正是 (恰如所期望的) $\mathbb{R}^{m+n}$ 上的 Lebesgue 测度.

设 $\mathcal{A}$ 和 μ 分别表示 $\mathbb{R}^n$ 上的 Lebesgue σ 代数和 Lebesgue 测度, 为简洁起见, $\mathcal{A}$ 的元简称为 $\mathbb{R}^n$ 中的可测子集, 简称 μ 为测度.

设 A 是 $\mathbb{R}^n$ 的可测子集. 如果某个性质在 A 上除了一个测度为零的可测子集的点外处处成立, 就称它在 A 上几乎处处 (a.e.) 成立, 或等价地称作对几乎所有的 $x\in A$ 成立. 例如, 称两个函数 $f,g:A\to\mathbb{R}$ 几乎处处相等是指集合 $\{x\in A; f(x)\neq g(x)\}$ 是可测的, 且测度为 0; 称函数 $f_n:A\to[-\infty,\infty]$ 的序列 $(f_n)_{n=1}^{\infty}$ 当 $n\to\infty$ 时在 A 中几乎处处收敛是指集合 $\{x\in A;\lim_{n\to\infty}f_n(x)$ 在 $[-\infty,\infty]$ 中存在 $\}$ 的余集是可测的且测度为零, 等等.

一个并不平凡的几乎处处成立的性质的例子是以下的基本结果. 下面在 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ 上赋以在第 1.10 节中定义的任何一个距离 d_p, $1\leqslant p\leqslant\infty$.

定理 1.14-2 (Rademacher 定理) [9] 设 Ω 为 $\mathbb{R}^n$ 的开子集, $f:\Omega\to\mathbb{R}^m$ 为 Lipschitz 连续函数 (1.11 节). 则 f 在 Ω 内几乎处处可微.

给定任意的 Lebesgue 可测子集 $A\in\mathcal{A}$ 和函数 $f:A\to[-\infty,\infty]$, 如果对任何 $\alpha\in\mathbb{R}$,

$$f^{-1}([-\infty,\alpha[)=\{x\in A; f(x)<\alpha\}\in\mathcal{A},$$

则称 f 为 Lebesgue 可测的, 或简称为可测的.

下面的定理列出了可测函数的一些基本性质. 注意, 性质 (c) 只适用于实值函数.

定理 1.14-3 设 A 为 $\mathbb{R}^n$ 为可测子集.

(a) 设 $f:A\to[-\infty,\infty]$ 为可测函数, 则函数 $|f|:A\to[0,\infty]$ 也是可测的.

---

[9] 这个命名是为了纪念 Hans Adolph Rademacher (1892—1969).

(b) 设函数 $f_n : A \to [-\infty, \infty]$, $n \geqslant 1$, 均为可测函数, 则函数

$$\sup_{n\geqslant 1} f_n, \quad \inf_{n\geqslant 1} f_n, \quad \limsup_{n\to\infty} f_n, \ \liminf_{n\to\infty} f_n : A \to [-\infty, \infty]$$

均是可测的.

(c) 设 $f, g : A \to \mathbb{R}$ 为可测函数, 则函数 $f + g : A \to \mathbb{R}$ 和函数 $fg : A \to \mathbb{R}$ 也是可测的.

下面的定理列出了另外三个基本性质, 它们将可测性与连续性联系起来. 性质 (c) 也称作 Lusin 性质.

定理 1.14-4 设 A 为 $\mathbb{R}^n$ 的可测子集.

(a) 设 $f : A \to \mathbb{R}$ 为连续函数, 则 f 是可测的.

(b) 设 $f : \mathbb{R} \to \mathbb{R}$ 为连续函数, $g : A \to \mathbb{R}$ 为可测函数, 则复合函数 $f \circ g : A \to \mathbb{R}$ 是可测的.

(c) 设 $f : \mathbb{R}^n \to \mathbb{R}$ 为可测函数, 使得 $\mu(A) < \infty$, 其中 $A := \{x \in \mathbb{R}^n; f(x) \neq 0\}$. 则对任何 $\varepsilon > 0$, 存在函数 $f_\varepsilon \in \mathcal{C}(\mathbb{R}^n)$, 其支集为 A 的紧子集, 满足

$$\sup_{x\in\mathbb{R}^n} |f_\varepsilon(x)| \leqslant \sup_{x\in\mathbb{R}^n} |f(x)| \text{ 且 } \mu(\{x \in \mathbb{R}^n; f(x) \neq f_\varepsilon(x)\}) \leqslant \varepsilon.$$

设 A 是 $\mathbb{R}^n$ 的任一可测子集. 所谓 A 上的一个简单函数是指函数 $s : A \to \mathbb{R}$, 其像是 $\mathbb{R}$ 的有限子集; 或等价地, 存在有限个 A 的互不相交的子集 A_i (即当 $i \neq j$ 时, $A_i \cap A_j = \varnothing$), $1 \leqslant i \leqslant m$ 和实数 α_i, $1 \leqslant i \leqslant m$, 使得

$$s = \sum_{i=1}^{m} \alpha_i \chi_{A_i},$$

其中 $\chi_{A_i} : A \to \mathbb{R}$ 表示集合 A_i 的特征函数 (1.2 节). 显然, 简单函数 s 为可测的充要条件是每个集合 A_i, $1 \leqslant i \leqslant m$, 是可测的.

下面的定理给出了关于可测性和可测简单函数的重要关联.

定理 1.14-5 设 A 为 $\mathbb{R}^n$ 的可测子集.

(a) 令 $f : A \to [-\infty, \infty]$ 为可测函数. 则存在一列可测的简单函数 $s_n : A \to \mathbb{R}$, $n \geqslant 1$, 使得对任何 $n \geqslant 1$, $|s_n| \leqslant |s_{n+1}| \leqslant |f|$ 且对任何 $x \in A$,

$$\lim_{n\to\infty} s_n(x) = f(x).$$

(b) $f : A \to [0, \infty]$ 为可测函数. 则存在一列可测的简单函数 $s_n : A \to \mathbb{R}$, $n \geqslant 1$, 使得对任何 $n \geqslant 1, 0 \leqslant s_n \leqslant s_{n+1} \leqslant f$ 且对任何 $x \in A$,

$$\lim_{n\to\infty} s_n(x) = f(x).$$

1.15 $\mathbb{R}^n$ 中的 Lebesgue 积分; 基本定理

设 A 为 $\mathbb{R}^n$ 的任一可测子集, 给定 A 上的一个可测的简单函数 $s = \sum_{i=1}^m \alpha_i \chi_{A_i}$ (1.14 节), 满足 $s \geqslant 0$ (等价于 $\alpha_i \geqslant 0,\ 1 \leqslant i \leqslant n$). 扩张实数 $\int_A s(x)dx \in [0, \infty]$ 定义为

$$\int_A s(x)dx := \sum_{i=1}^m \alpha_i \mu(A_i).$$

任意的可测函数 $f : A \to [0, \infty]$ 的 Lebesgue 积分定义为

$$\int_A f(x)dx := \sup \left\{ \int_A s(x)dx; s \text{ 为可测的简单函数, 在 } A \text{ 上 } 0 \leqslant s \leqslant f \right\}.$$

因此, $\int_A f(x)dx$ 或者 $\geqslant 0$, 或者等于 ∞. 如果可测函数 $f : A \to [0, \infty]$ 的 Lebesgue 积分 $\int_A f(x)dx$ 是有限数, 则 f 必定几乎处处有限.

最后, 函数 $f : A \to [-\infty, \infty]$ 称为 Lebesgue 可积的, 或简称为可积的, 是指它是可测的, 满足

$$\int_A \max\{f(x); 0\}dx < \infty \text{ 且 } \int_A \max\{-f(x); 0\}dx < \infty.$$

如果 $f : A \to [-\infty, \infty]$ 是 Lebesgue 可积的, 其 Lebesgue 积分定义为

$$\int_A f(x)dx := \int_A \max\{f(x); 0\}dx - \int_A \max\{-f(x); 0\}dx.$$

经常要用到下列关于 Lebesgue 可积函数及其 Lebesgue 积分的性质: 可积函数的 Lebesgue 积分是一个实数, 即它不等于 ∞ 或 $-\infty$; $\mathbb{R}^n$ 中可测子集的 Lebesgue 测度也可由 $\int_{\mathbb{R}^n} \chi_A dx = \int_A dx$ 给出; Lebesgue 可积函数是几乎处处有限的; 可测函数 $f : A \to [-\infty, \infty]$ 是 Lebesgue 可积的当且仅当函数 $|f| : A \to [0, \infty]$ 是 Lebesgue 可积的, 此时

$$\left| \int_A f(x)dx \right| \leqslant \int_A |f(x)|dx.$$

所有 Lebesgue 可积函数 $f : A \to \mathbb{R}$ 的集合记作

$$\mathcal{L}^1(A),$$

它显然是 $\mathbb{R}$ 上的向量空间 ($\mathbb{R}$ 或 $\mathbb{C}$ 上的向量空间定义于 2.1 节).

在空间 $\mathcal{L}^1(A)$ 中, 由关系 "在 A 上 $f = g$ a.e." 定义了等价关系 $\mathcal{R}$, 商集

$$L^1(A) := \mathcal{L}^1(A)/\mathcal{R}$$

也是 $\mathbb{R}$ 上的向量空间. 此外, 如果 $f, g \in \mathcal{L}^1(A)$, 且在 A 上 $f = g$ a.e., 则

$$\int_A f(x)dx = \int_A g(x)dx.$$

由此, $L^1(A)$ 中任何一个等价类的 Lebesgue 积分可以一意地定义为这个类中任何一个函数的 Lebesgue 积分.

通常我们称 $L^1(A)$ 中的元为可积函数, 尽管实际上它们是可积函数模 $\mathcal{R}$ 的等价类.

显然, 把 $\mathcal{L}^1(A)$ 中的函数等同于它在 $L^1(A)$ 中的等价类, 在语言上难避违规滥用之嫌, 但它却可免去许多累赘的叙述, 而且也不会引起任何误解. 例如, "$f \in L^1(A)$ 为连续函数" 意为在 f 的等价类中, 有一个 (唯一的) 在 $\mathcal{L}^1(A)$ 中的连续函数; 类似地, "$f \in L^1(A)$ 在 A 中几乎处处有限" 意为在 f 的等价类中, 有一个在 A 上处处有限的函数, 等等.

可以通过其他途径定义 Lebesgue 积分和空间 $L^1(A)$. 例如, 令 $\mathcal{S}(A)$ 表示由所有可积的简单函数 $s : A \to \mathcal{R}$, 即满足

$$\mu(\{x \in A; s(x) \neq 0\}) < \infty$$

的简单函数组成的集合. 易见商集 $S(A) := \mathcal{S}(A)/\mathcal{R}$ 为一个向量空间, 映射

$$\|\cdot\|_{L^1(A)} : s \in S(A) \to \int_A |s(x)|dx,$$

其中 A 上简单函数的积分如前所定义, 是 $S(A)$ 上的范数.

这样, 空间 $L^1(A)$ 可以定义为空间 $(S(A), \|\cdot\|_{L^1(A)})$ 的完备化空间 (定理 1.12-4). 此时, 函数 $f \in L^1(A)$ 的 Lebesgue 积分是定义在 $L^1(A)$ 的稠密子集 $S(A)$ 上的连续线性泛函

$$s \in S(A) \to \int_A s(x)dx$$

的唯一的连续延拓 (定理 1.12-3).

当 A 是 $\mathbb{R}^n$ 的开子集时, 还可给出另一个定义: 用 $\mathcal{C}_c(A)$ 表示在 A 上所有具有紧支集的连续函数 $f : A \to \mathbb{R}$ 组成的空间, 令

$$\|f\|_{L^1(A)} = \int_A |f(x)|dx,$$

其中记号 $\int_A g(x)dx$ 表示函数 $g \in \mathcal{C}_c(A)$ 的 Riemann 积分. 于是, 空间 $L^1(A)$ 也可等价地定义为空间 $(\mathcal{C}_c(A), \|\cdot\|_{L^1(A)})$ 的完备化空间, 由构造, $\mathcal{C}_c(A)$ 在 $L^1(A)$ 中稠密 (若 $L^1(A)$ 如本节前面所定义, 则 $\mathcal{C}_c(A)$ 在 $L^1(A)$ 中的稠密性是一个需要证明的定理, 见定理 2.5-3). 此时, Lebesgue 积分也定义为在一个稠密子集上的连续线性泛函的唯一连续延拓.

Lebesgue 可积函数的概念很容易推广到复值函数上: 设 A 为 $\mathbb{R}^n$ 的任意可测子集, 复值函数 $f : A \to \mathbb{C}$ 称为 Lebesgue 可积的, 是指

$$\mathrm{Re}\, f \in \mathcal{L}^1(A) \text{ 且 } \mathrm{Im}\, f \in \mathcal{L}^1(A),$$

此时, f 的 Lebesgue 积分定义为

$$\int_A f(x)dx := \int_A \mathrm{Re}\, f(x)dx + i\int_A \mathrm{Im}\, f(x)dx.$$

容易知道它仍满足不等式

$$\left|\int_A f(x)dx\right| \leqslant \int_A |f(x)|dx.$$

所以 Lebesgue 可积函数 $f: A \to \mathbb{C}$ 的集合记作

$$\mathcal{L}^1(A;\mathbb{C}),$$

它显然是一个 $\mathbb{C}$ 上的向量空间. 关系 "在 A 中 $f = g$, a.e." 也在空间 $\mathcal{L}^1(A;\mathbb{C})$ 上定义了一个等价关系 $\mathcal{R}$, 商集

$$L^1(A;\mathbb{C}) := \mathcal{L}^1(A;\mathbb{C})/\mathcal{R}$$

也是一个 $\mathbb{C}$ 上的向量空间.

最后要注意的是, 为简明起见, 如果 $f \in L^1(A)$ 或 $f \in L^1(A;\mathbb{C})$, 我们经常略去变量 $x \in A$, 简记作

$$\int_A f dx := \int_A f(x)dx.$$

下面的定理给出了 Lebesgue 积分的最基本的性质. 注意, 在这里建立这些性质的次序实际上取决于前述定义 Lebesgue 积分的方式. 前面三个定理列出了可积函数序列的三个基本的收敛性质.

定理 1.15-1 (Beppo Levi 单调收敛定理) 设 A 为 $\mathbb{R}^n$ 中的可测子集, $(f_k)_{k=1}^\infty$ 是函数 $f_k \in \mathcal{L}^1(A)$ 的序列, 满足

$$0 \leqslant f_1 \leqslant f_2 \leqslant \cdots \leqslant f_k \leqslant f_{k+1} \leqslant \cdots, \text{ 在 } A \text{ 上 a.e. } \lim_{k\to\infty}\int_A f_k(x)dx < \infty.$$

则存在函数 $f \in \mathcal{L}^1(A)$, 使得

$$\lim_{k\to\infty} f_k(x) = f(x), \text{ 在 } A \text{ 上 a.e. } \lim_{k\to\infty}\int_A |f_k(x) - f(x)|dx = 0.$$

特别地, 对几乎所有的 $x \in A$, $\lim_{k\to\infty} f_k(x) < \infty$.

定理 1.15-2 (Fatou 引理) 设 A 为 $\mathbb{R}^n$ 为可测子集, $(f_k)_{k=1}^\infty$ 为可测函数 $f_k: A \to \mathbb{R}$ 的序列, 满足在 A 上

$$f_k \geqslant 0 \quad \text{a.e.}$$

则

$$\int_A (\liminf_{k\to\infty} f_k(x))dx \leqslant \lim_{k\to\infty}\inf \int_A f_k(x)dx,$$

这个不等式的右边或两边允许等于 ∞.

定理 1.15-3 (Lebesgue 控制收敛定理) 设 A 为 $\mathbb{R}^n$ 的可测子集, $(f_k)_{k=1}^{\infty}$ 为函数 $f_k \in \mathcal{L}^1(A)$ 或相应的 $f_k \in \mathcal{L}^1(A;\mathbb{C})$; 使得对几乎所有的 $x \in A$, 存在

$$f(x) := \lim_{k\to\infty} f_k(x),$$

而且存在函数 $g \in \mathcal{L}^1(A)$, 使得对所有的 $k \geqslant 1$,

$$|f_k(x)| \leqslant g(x), \quad \text{a.e.}$$

则 $f \in \mathcal{L}^1(A)$, 或相应地 $f \in \mathcal{L}^1(A;\mathbb{C})$, 而且

$$\lim_{k\to\infty} \int_A |f_k(x) - f(x)| dx = 0.$$

特别地,

$$\int_A f(x)dx = \lim_{k\to\infty} \int_A f_k(x)dx.$$

设 B 为 $\mathbb{R}^n$ 的可测子集, $\mathcal{A}$ 为 B 的所有 Lebesgue 可测子集组成的 σ 代数. 给定函数 $f \in \mathcal{L}^1(B)$, 对所有的 $A \in \mathcal{A}$, 令

$$\nu(A) := \int_A f(x)dx.$$

于是, 对每个 $A \in \mathcal{A}$, $|\nu(A)| \leqslant \int_A |f(x)|dx \leqslant \int_B |f(x)|dx < \infty$. 显然, 这样定义的函数 $\nu : \mathcal{A} \to \mathbb{R}$ 具有下列性质: 首先, 它是一个带号测度, 即

$$\nu(\varnothing) = 0.$$

如果 $A_i \in \mathcal{A}$, $i \geqslant 1$ 且当 $i \neq j$ 时, $A_i \cap A_j = \varnothing$, 则 $\nu(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \nu(A_i)$ (可列可加性易由 Lebesgue 控制收敛定理导出). 其次, 它关于 Lebesgue 测度 dx 是绝对连续的, 即

$$A \in \mathcal{A} \quad \text{且当} \quad dx\text{-meas } A = 0 \text{ 时 } \nu(A) = 0.$$

值得注意的是反向的性质成立:

定理 1.15-4 (Radon-Nikodym 定理) 设 B 为 $\mathbb{R}^n$ 的可测子集, $\mathcal{A}$ 为 B 的所有可测子集组成的 σ 代数, $\nu : \mathcal{A} \to \mathbb{R}$ 为带号测度, 且关于 Lebesgue 测度绝对连续, 则存在函数 $f \in \mathcal{L}^1(B)$, 使得对每个 $A \in \mathcal{A}$,

$$\nu(A) = \int_A f(x)dx.$$

下面的定理给出了定义在 $\mathbb{R}^m \times \mathbb{R}^n$ 中可测集乘积集合上的函数的 Lebesgue 可积性的基本特征, 同时, 也给出了计算 Lebesgue 可积函数 $f : A \times B \to [-\infty, \infty]$ 的 Lebesgue 积分

$$\iint_{A\times B} f(x,y)dxdy,$$

的方法, 其中 $dxdy$ 表示 $\mathbb{R}^m \times \mathbb{R}^n$ 上的 Lebesgue 测度.

定理 1.15-5 设 A 为 $\mathbb{R}^m$ 的可测子集, B 为 $\mathbb{R}^n$ 的可测子集, $f: A\times B\subset\mathbb{R}^m\times\mathbb{R}^n\to[-\infty,\infty]$ 为可测函数.

(a) (Tonelli 定理) 对每个 $x\in A$, 函数 $f(x,\cdot): y\in B\to f(x,y)\in[-\infty,\infty]$ 是可测的, 对每个 $y\in B$, 函数 $f(\cdot,y): x\in A\to f(x,y)\in[-\infty,\infty]$ 是可测的. 而且, 函数 f 在 $A\times B$ 上可积, 即

$$\iint_{A\times B}|f(x,y)|dxdy<\infty$$

当且仅当下列两条件之一满足:

$$\int_A\left(\int_B|f(x,y)|dy\right)dx<\infty,$$
$$\int_B\left(\int_A|f(x,y)|dx\right)dy<\infty.$$

(b) (Fubini 定理) 如果函数 f 在 $A\times B$ 上可积, 则函数 $f(\cdot,y): A\to[-\infty,\infty]$ 对几乎所有的 $y\in B$ 可积, 函数 $f(x,\cdot): B\to[-\infty,\infty]$ 对几乎所有的 $x\in A$ 可积, f 在 $A\times B$ 上的 Lebesgue 积分由

$$\iint_{A\times B}f(x,y)dxdy=\int_A\left(\int_B f(x,y)dy\right)dx=\int_B\left(\int_A f(x,y)dx\right)dy$$

给出.

1.16 $\mathbb{R}^n$ 上 Lebesgue 积分的变量代换

在本节中, 我们来考察定义于 $\mathbb{R}^n$ 开子集上的 Lebesgue 积分如何进行变量代换, 即这里的开集是 $\mathbb{R}^n$ 中另一开子集 Ω 在映射 $\varphi=(\varphi_i)_{i=1}^n:\Omega\to\mathbb{R}^n$ 下的像 $\varphi(\Omega)$, 且在此过程中变量 $y\in\varphi(\Omega)$ 被变量 $x\in\Omega$ 所取代.

记号 $\nabla\varphi$ 表示 $n\times n$ 矩阵域, 定义为 $(\nabla\varphi)_{ij}=\partial_j\varphi_i$, $1\leqslant i,j\leqslant n$, 其中 ∂_j 表示对第 j 个变量的偏导算子.

定理 1.16-1 ($\mathbb{R}^n$ 上 Lebesgue 积分的变量单射变换) 设 Ω 为 $\mathbb{R}^n$ 的开子集, $\varphi:\Omega\to\mathbb{R}^n$ 为连续可微的单射.

则函数 $f:\varphi(\Omega)\to\mathbb{R}$ 在 $\varphi(\Omega)$ 上 Lebesgue 可积的充要条件是函数

$$x\in\Omega\to f(\varphi(x))|\det\nabla\varphi(x)|\in\mathbb{R}$$

在 Ω 上 Lebesgue 可积. 此时,

$$\int_{\varphi(\Omega)}f(y)dy=\int_\Omega f(\varphi(x))|\det\nabla\varphi(x)|dx.$$

注 (1) 在定理 1.16-1 的假设下, 集合 $\varphi(\Omega)$ 自动成为开集 (因而 $\varphi(\Omega)$ 上的 Lebesgue 可积性有意义): 这是深刻的 $\mathbb{R}^n$ 中区域的 Brouwer 不变性定理 (定理 9.17-3) 的推论, 实际上只需假定 $\varphi:\Omega\to\mathbb{R}$ 连续即可.

(2) 因为 Lebesgue 可积函数是几乎处处有限的, f 取实值并非限制性的假设.

在许多教材中处理的都是映射 φ 为单射的情况 (定理 1.16-1), 而 φ 非单射的情况 (如下面的定理) 并不常见[10].

定理 1.16-2 ($\mathbb{R}^n$ 上 Lebesgue 积分的变量非单射变换) 设 Ω 为 $\mathbb{R}^n$ 的开子集, $\varphi:\Omega\to\mathbb{R}^n$ 为连续可微映射, 像 $\varphi(\Omega)$ 为开集. 对每个 $y\in\varphi(\Omega)$, 令

$$\mathrm{card}\,\varphi^{-1}(y) = \text{集 } \varphi^{-1}(y) \text{ 的基数, 如果 } \varphi^{-1}(y) \text{ 为有限集,}$$
$$\mathrm{card}\,\varphi^{-1}(y) = \infty, \text{ 如果 } \varphi^{-1}(y) \text{ 为无限集.}$$

则对给定的函数 $f:\varphi(\Omega)\to\mathbb{R}$, 函数 $f\mathrm{card}\,\varphi^{-1}:\varphi(\Omega)\to\mathbb{R}$ 在 $\varphi(\Omega)$ 上 Lebesgue 可积的充要条件是函数

$$x\in\Omega\to f(\varphi(x))|\det\nabla\varphi(x)|dx\in\mathbb{R}$$

在 Ω 上 Lebesgue 可积. 如果上述条件成立, 那么

$$\int_{\varphi(\Omega)} f(y)\mathrm{card}\,\varphi^{-1}(y)dy = \int_\Omega f(\varphi(x))|\det\nabla\varphi(x)|dx.$$

注 (1) 如所期望的, 定理 1.16-1 是定理 1.16-2 的特殊情况.

(2) 与定理 1.16-1 不同的是定理 1.16-2 中必须假设 $\varphi(\Omega)$ 是开集.

1.17 $\mathbb{R}^n$ 中的体积、面积和长度

$\mathbb{R}^n$ 中可测子集 A 的 n 维体积或简称为体积, 记作 dx-meas A 或简记作 meas A, 即定义为 A 的 Lebesgue 测度, 换句话说,

$$dx\text{-meas } A = \text{meas } A := \int_A dx,$$

这里 dx 表示 n 维 Lebesgue 测度.

利用重积分的变量代换公式 (定理 1.16-1 或定理 1.16-2), 我们可以计算 n 维平行多面体 ($\mathbb{R}^n$ 中这类特殊的子集的定义见下面的定理) 的体积.

[10] 见 Rado & Reichelderfer [1955], Schwartz [1993b, 推论 6.2.14], Federer [1969] 或 Smith [1983, 第 16 章].

定理 1.17-1 **(n 维平行多面体的体积)** 设 P 为 $\mathbb{R}^n$ 中 n 维平行多面体, 即 $\mathbb{R}^n$ 中形为

$$P = \left\{ a + \sum_{i=1}^{n} \lambda_i b_i; 0 \leqslant \lambda_i \leqslant 1, 1 \leqslant i \leqslant n \right\},$$

其中 $a \in \mathbb{R}^n$ 且 $b_i \in \mathbb{R}^n$, $1 \leqslant i \leqslant n$. 则 P 的体积为

$$dx\text{-meas}\, P = |\det B| = \sqrt{\det(b_i \cdot b_j)},$$

其中 B 为 $n \times n$ 矩阵, 其第 i 列即向量 b_i (这里可视为 $n \times 1$ 矩阵), $(b_i \cdot b_j)$ 表示 $n \times n$ 矩阵, 其第 i 行, 第 j 列的元即向量 b_i 和 b_j 的 Euclid 内积.

注 (1) dx-meas P 的第二个表达式是第一个表达式的直接推论 (因为对任何方阵 $B, (\det B)^2 = \det(B^\top B)$).

(2) 如果向量 b_i 线性相关, 则平行多面体 P 的体积为零.

注意到在 $\mathbb{R}^n$ 的直交标架变换时, 上述矩阵 B 的元一般也会变化, 但由于直交标架变换时 Euclid 内积不变, 因而矩阵 $(b_i \cdot b_j)$ 不变. 因此, 可以用第二个表达式来定义作为 $\mathbb{R}^m$ 的子集的 n 维平行多面体的 n 维体积, 这里 $m \geqslant n$. 这一点也是下面关于 n 维面积的定义的基础.

设 Ω 是 $\mathbb{R}^n$ 的开子集, $m \geqslant n$, $\Theta = (\Theta_j)_{j=1}^m : \Omega \to \mathbb{R}^m$ 为连续可微的单射, 在每个点 $x \in \Omega$, 矩阵 $(\nabla\Theta)(x) \in \mathbb{M}^{m\times n}$, 其中 $(\nabla\Theta)_{ij} = \partial_j \Theta_i$, $(\nabla\Theta)(x)$ 映 $\mathbb{R}^n$ 中 n 个基向量为 n 个向量 $\partial_i\Theta(x) := (\partial_i\Theta_j(x))_{j=1}^m \in \mathbb{R}^m$, $1 \leqslant i \leqslant n$, 由它们又定义了 $\mathbb{R}^m$ 中形如

$$\left\{ \Theta(x) + \sum_{i=1}^{n} \lambda_i \partial_i \Theta(x);\ 0 \leqslant \lambda_i \leqslant 1,\ 1 \leqslant i \leqslant n \right\}$$

的 n 维平行多面体.

因为由定理 1.17-1, 这个平行多面体的 n 维体积为

$$\sqrt{\det(\partial_i\Theta(x) \cdot \partial_j\Theta(x))},$$

于是, 可以自然地定义集合 $\Theta(\Omega)$ 的 n 维面积, 或简称面积, area $\Theta(\Omega)$ 即称为 "基本 n 维体积 $\sqrt{\det(\partial_i\Theta(x) \cdot \partial_j\Theta(x))}dx$ 的无限和", 即

$$\text{area}\, \Theta(\Omega) := \int_\Omega \sqrt{\det(\partial_i\Theta(x) \cdot \partial_j\Theta(x))}dx.$$

注 (1) 如果 $m = n$, 且 $\Theta = \text{id}_\Omega$, 则 $\Theta(\Omega) = \Omega$ 的面积 (如所设想的) 即上面定义的 Ω 的 n 维体积.

(2) 如果 $m = n$, 而且 Θ 是浸入, 即对每个 $x \in \Omega$, 矩阵 $(\nabla\Theta)(x)$ 可逆, 则矩阵 $(\partial_i\Theta(x) \cdot \partial_j\Theta(x)) \in \mathbb{M}^n$ 是集合 $\Theta(\Omega)$ 在 $x \in \Omega$ 处的距离张量, 见第 8.2 节.

(3) 如果 $n=2, m=3$, 而且 Θ 是浸入, 即在每一点 $x \in \Omega$, 矩阵 $(\partial_i\Theta(x)\cdot\partial_j\Theta(x))$ 秩为 2, 矩阵 $(\partial_i\Theta(x)\cdot\partial_j\Theta(x)) \in \mathbb{M}^2$ 是集合 $\Theta(\Omega)$ 在 $x \in \Omega$ 处的第一基本型, 因而称 $\Theta(\Omega)$ 为 $\mathbb{R}^3$ 中的曲面, 见第 8.9 节.

最后, 设 $n=1$, 即考察集合 Ω 是 $\mathbb{R}$ 中的开区间, $\Theta=(\Theta_j)_{j=1}^m$ 是 I 到 $\mathbb{R}^m$ 的单射, 这里 $m \geqslant 1$. 此时, 称区间 I 在 Θ 下的像为 $\mathbb{R}^m$ 中的曲线, 变量 $t \in I$ 称为曲线 $\Theta(I)$ 的参数. 曲线 $\Theta(I)$ 的长度自然定义为集合 $\Theta(I)$ 的一维面积, 即

$$\Theta(I) \text{ 的长度} := \int_I \sqrt{\Theta'(t)\cdot\Theta'(t)}dt,$$

其中 $\Theta'(t)=(\Theta_j'(t))_{j=1}^m \in \mathbb{R}^m, t \in I$, 注意, 被积函数 $\sqrt{\Theta'(t)\cdot\Theta'(t)}$ 即向量 $\Theta'(t) \in \mathbb{R}^m$ 的 Euclid 范数, 见第 2.2 节.

如果对所有的 $t \in I, \Theta'(t) \neq 0, t_0 \in I$, 则以 $\Theta(t_0)$ 为起点沿着曲线 $\Theta(I)$ 的弧长定义为

$$s := \sigma(t) = \int_{t_0}^t \sqrt{\Theta'(t)\cdot\Theta'(t)}dt,$$

这样定义的函数 $\sigma: I \to \mathbb{R}$ 是可逆的, 其反函数 $\tau: \sigma(I) \to I$ 在所有的 $s=\sigma(t)$, $t \in I$ 处的导数由

$$\tau'(s) = \frac{1}{\sqrt{\Theta'(t)\cdot\Theta'(t)}}$$

给出.

1.18 空间 $\mathcal{C}^m(\Omega)$ 和 $\mathcal{C}^m(\overline{\Omega})$; $\mathbb{R}^n$ 中的域

本书中考虑的所有函数均为实值的.

点 $x \in \mathbb{R}^n$ 的坐标记作 x_i, $1 \leqslant i \leqslant n$, 相应的偏导算子记作 $\partial_i := \partial/\partial x_i$, $\partial_{ij} := \partial^2/\partial x_i\partial x_j$, $\partial_{ijk} := \partial^3/\partial x_i\partial x_j\partial x_k$ 等. 任意阶的偏导算子也用多重指标记作

$$\partial^{\boldsymbol{\alpha}} := \partial^{|\boldsymbol{\alpha}|}/\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\cdots\partial x_n^{\alpha_n},$$

其中 $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_n), \alpha_i \in \mathbb{N}$, $1 \leqslant i \leqslant n$, 为多重指标, $|\boldsymbol{\alpha}| := \sum_{i=1}^n \alpha_i \geqslant 0$, 注意, 允许 $\mathbf{0}=(0,0,\cdots,0)$, 并约定 $\alpha^{\mathbf{0}}v := v$. 最后, 若 $\boldsymbol{x}=(x_i) \in \mathbb{R}^n$, 令 $|\boldsymbol{x}| := \left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}}$ (函数 $|\cdot|: \mathbb{R}^n \to \mathbb{R}$ 按此定义为 Euclid 范数, 见 2.2 节).

首先, 设 Ω 为 $\mathbb{R}^n$ 的任意开子集 (本节后面将会增添对集合 Ω 的假设).

对任何整数 $m \geqslant 1$, Ω 上 m 阶和无限阶连续可微函数组成的空间分别记作

$$\mathcal{C}^m(\Omega) \quad \text{和} \quad \mathcal{C}^\infty(\Omega) := \bigcap_{m=0}^{\infty} \mathcal{C}^m(\Omega),$$

对 $m=0$, 令

$$\mathcal{C}^0(\Omega) := \mathcal{C}(\Omega), \quad \mathcal{C}^0(\overline{\Omega}) := \mathcal{C}(\overline{\Omega}).$$

对每个整数 $m \geqslant 1$, 定义空间

$$\mathcal{C}^m(\overline{\Omega}) := \{f \in \mathcal{C}^m(\Omega);\ 对每个\ |\boldsymbol{\alpha}| \leqslant m,\ 存在\ g^{\boldsymbol{\alpha}} \in \mathcal{C}(\overline{\Omega}),\ 使得\ \partial^{\boldsymbol{\alpha}} f = g^{\boldsymbol{\alpha}}|_{\Omega}\}.$$

换言之, $\mathcal{C}^m(\overline{\Omega})$ 是由连同其各个偏导数 $\partial^{\boldsymbol{\alpha}} f$, $1 \leqslant |\boldsymbol{\alpha}| \leqslant m$ 均可连续延拓到 $\overline{\Omega}$ 上的所有函数 $f \in \mathcal{C}^m(\Omega)$ 组成, 这又等价于由 $\mathcal{C}^m(\Omega)$ 中满足对任何 $x_0 \in \partial\Omega$, $0 \leqslant |\boldsymbol{\alpha}| \leqslant m$, $\lim_{x\to x_0} \partial^{\boldsymbol{\alpha}} f(x)$ 在 $\mathbb{R}$ 中存在的函数组成. 当 Ω 有界时, 也等价于由使 $\partial^{\boldsymbol{\alpha}} f$, $0 \leqslant |\boldsymbol{\alpha}| \leqslant m$, 在 Ω 一致连续的函数组成.

用 $\mathcal{C}^{m,\lambda}(\overline{\Omega})$ 表示 $\mathcal{C}^m(\overline{\Omega})$ 的一个子空间, 当 $0 < \lambda < 1$ 时由所有 m 阶偏导数满足指数为 λ 的 Hölder 条件的函数组成, 当 $\lambda = 1$ 时, 由所有 m 阶偏导数满足 Lipschitz 连续性的函数组成 (1.11 节), 即

$$\begin{aligned}\mathcal{C}^{m,\lambda}(\overline{\Omega}) : = \{f \in \mathcal{C}^m(\overline{\Omega});\ &存在数\ L,\ 对所有\ |\boldsymbol{\alpha}| = m\ 和\\ &x, y \in \overline{\Omega},\ 成立\ |\partial^{\boldsymbol{\alpha}} f(x) - \partial^{\boldsymbol{\alpha}} f(y)| \leqslant L|x-y|^{\lambda}\}.\end{aligned}$$

$\mathbb{R}^n$ 中开子集 Ω 的边界 Γ 称为 Lipschitz 连续的, 是指它满足下列条件: 即存在常数 $\alpha > 0$, $L > 0$ 和有限个局部坐标系, 其坐标 $\boldsymbol{\zeta}_r' = (\zeta_1^r, \zeta_2^r, \cdots, \zeta_{n-1}^r) \in \mathbb{R}^{n-1}, \zeta_r = \zeta_n^r$, 及相应的函数 $\theta_r : \omega_r := \{\boldsymbol{\zeta}_r \in \mathbb{R}^{n-1}; |\boldsymbol{\zeta}_r| < \alpha\} \to \mathbb{R}$, $1 \leqslant r \leqslant s$, 使得

$$\Gamma = \bigcup_{r=1}^{s} \{(\boldsymbol{\zeta}_r', \zeta_r); \boldsymbol{\zeta}_r' \in \omega_r\ 且\ \zeta_r = \theta_r(\boldsymbol{\zeta}_r')\},$$

而且对所有的 $\boldsymbol{\zeta}_r', \boldsymbol{\eta}_r' \in \omega_r$, $1 \leqslant r \leqslant s$, 有

$$|\theta_r(\zeta_r') - \theta_r(\eta_r')| \leqslant L|\zeta_r' - \eta_r'|.$$

后一不等式表示映射 θ_r 满足 Lipschitz 连续性. 注意, 为记号使用方便, $\{(\boldsymbol{\zeta}_r', \zeta_r); \boldsymbol{\zeta}_r' \in \omega_r, \zeta_r = \theta_r(\boldsymbol{\zeta}_r')\}$ 表示在第 r 个局部坐标系中坐标 ζ_i^r, $1 \leqslant i \leqslant n$ 满足 $|(\zeta_1^r, \zeta_2^r, \cdots, \zeta_{n-1}^r)| < \alpha$, 且 $\zeta_n^r = \theta_r(\zeta_1^r, \zeta_2^r, \cdots, \zeta_{n-1}^r)$ 的点组成的集合.

注 虽然, 具有 Lipschitz 连续性的边界 Γ 必定有界, 而集合 Ω 却未必有界, 在定义中它可用集合 $\mathbb{R}^n - \overline{\Omega}$ 取代.

类似地, 如果映射 θ_r, $1 \leqslant r \leqslant s$, 属于空间 $\mathcal{C}^m(\omega_r)$, 其中 $m \geqslant 1$, 则称边界 Γ 属于 $\mathcal{C}^m$ 类.

更一般地, Γ 的子集 Γ_0 称为 Lipschitz 连续的或 $\mathcal{C}^m$ 类的, 是指在定义中用 Γ_0 代替 Γ.

称开集 Ω *局部地在边界 Γ 的同一侧*, 是指存在常数 $\beta > 0$, 使得

$$\begin{aligned}&\{(\zeta_r', \zeta_r);\ \zeta_r' \in \omega_r,\ 且\ \theta_r(\zeta_r') < \zeta_r < \theta_r(\zeta_r') + \beta\} \subset \Omega,\ 1 \leqslant r \leqslant s,\\ &\{(\zeta_r', \zeta_r);\ \zeta_r' \in \omega_r,\ 且\ \theta_r(\zeta_r') - \beta < \zeta_r < \theta_r(\zeta_r')\} \subset \mathbb{R}^n - \overline{\Omega},\ 1 \leqslant r \leqslant s.\end{aligned}$$

图 1.18-1 $\mathbb{R}^2$ 中的域. 该图原见于 P. G. CIARLET [1988]: Mathematical Elasticity, Volume I: Three-Dimensional Elasticity. North-Holland, Amsterdam

图 1.18-2 $\Omega \subset \mathbb{R}^2$ 中有界开子集并非为域的例子. 该图原见于 P. G. CIARLET [1988]: Mathematical Elasticity, Volume I: Three-Dimensional Elasticity. North-Holland, Amsterdam

所谓 $\mathbb{R}^n$ 中的域 Ω 是指 $\mathbb{R}^n$ 中边界 Γ 具有 Lipschitz 连续性的有界连通开子集 Ω, 其中 Ω 局部地在 Γ 的同一侧 (当 $n = 2$ 时见图 1.18-1 和反例见图 1.18-2).

当 Ω 是一个域 (替代本节此前的约定为任意开子集), 由于域的一个关键性质, 可以给出空间 $\mathcal{C}^m(\overline{\Omega})$ 的另一个等价定义 (定理 1.18-1). 注意, 下面的定理[11] 可以视作

[11] 其证明例如可见 Stein [1970, 第 6 章], 或者:

P. G. Ciarlet; C. Mardare [2004]: Recovery of a manifold with boundary and its continuity as a function of its metric tensor. Journal de Mathématiques Pures et Appliquées 83, 811–843.

如可预见的, 其证明复杂, 特别地, 它依赖深刻的扩张定理, 见:

H. Whitney [1934]: Analytic extensions of differentiable functions defined in closed sets, Transactions of the American Mathematical Society 36, 63–89.

连续函数 ($m=0$) 的 Tietz-Urysohn 扩张定理 (定理 1.7-7) 在现在的情况下推广到连续可微函数上 ($m \geqslant 1$).

定理 1.18-1 设 Ω 为 $\mathbb{R}^n$ 中的一个域, 则对任何整数 $m \geqslant 1$ 和 $m=\infty$, 空间 $\mathcal{C}^m(\overline{\Omega})$ 也可以定义为

$$\mathcal{C}^m(\overline{\Omega}) = \{f|_\Omega; f \in \mathcal{C}^m(\mathbb{R}^n)\}.$$

边界 Lipschitz 连续性的意义在于它们虽然并不充分光滑, 但正如此处要简略说明的, 可以定义其上的曲面积分, 而且 Green 公式成立. 这里并不讨论所涉函数的可测性.

函数 $f:\Gamma \to \mathbb{R}$ 称为 $d\Gamma$ 几乎处处可定义的, 是指每个函数 $\zeta_r' \in \omega_r \to f(\zeta_r', \theta_r(\zeta_r'))$, $1 \leqslant r \leqslant s$, 在集合 ω_r 上几乎处处有定义 (按 $n-1$ 维 Lebesgue 测度). 如果每个函数 $\zeta_r' \in \omega_r \to f(\zeta_r', \theta_r(\zeta_r'))$ 还是 Lebesgue 可积的, 即

$$\int_{\omega_r} |f(\zeta_r', \theta_r(\zeta_r'))| d\zeta_r' < \infty,$$

则称函数 f 在 Γ 上可积, 所有这类函数组成的向量空间记作 $\mathcal{L}^1(\Gamma)$.

为定义函数 $f \in \mathcal{L}^1(\Gamma)$ 的积分, 我们需要和边界 Γ 的与开集 U_r 的覆盖相应的单位分解 (图 1.18-3), 其中

$$U_r := \{(\zeta_r', \zeta_r)\};\ \zeta_r' \in \omega_r \text{ 且 } \theta_r(\zeta_r') - \beta < \zeta_r < \theta_r(\zeta_r') + \beta\},$$

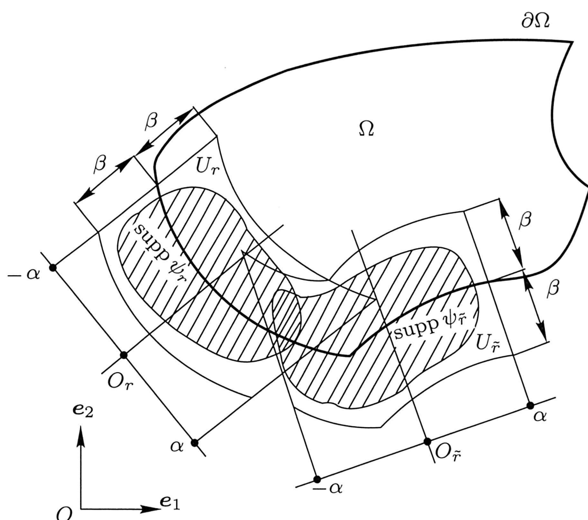


图 1.18-3 $\mathbb{R}^2$ 中域的边界 Γ 覆盖 $\Gamma \subset \bigcup_{r=1}^s U_r$ 相应的单位分解中, 两个函数 $\psi_r, \psi_{\tilde{r}}$ 的交集. 该图原见于 P. G. Ciarlet [1988]: Mathematical Elasticity, Volume I: Three-Dimensional Elasticity. North-Holland, Amsterdam

即函数族 $\psi_r \in \mathcal{C}^\infty(\mathbb{R}^n)$, $1 \leqslant r \leqslant s$, 满足

$$\operatorname{supp} \psi_r \subset U_r, \quad 0 \leqslant \psi_r \leqslant 1, \quad 1 \leqslant r \leqslant s,$$
$$\sum_{r=1}^{s} \psi_r(x) = 1, \quad \text{对所有 } x \in \Gamma.$$

函数 $f \in \mathcal{L}^1(\Gamma)$ 的曲面积分定义为

$$\int_\Gamma f d\Gamma := \sum_{r=1}^{s} \int_{\omega_r} f(\zeta_r', \theta_r(\zeta_r'))\psi_r(\zeta_r', \theta_r(\zeta_r')) \left(1 + \sum_{i=1}^{n-1} \left|\frac{\partial \theta_r}{\partial \zeta_i^r}\right|^2\right)^{\frac{1}{2}} d\zeta_r',$$

称 $d\Gamma$ 为沿 Γ 的面积元素. 这个定义是有意义的. 首先, 由 Rademacher 定理 (定理 1.14-2), 因为函数 θ_r 为 Lipschitz 连续的, 所以它 (按 $n-1$ 维 Lebesgue 测度) 几乎处处可微, 且其偏导数满足对几乎所有的 $\zeta_r' \in \omega_r$ 和 $1 \leqslant i \leqslant n-1, 1 \leqslant r \leqslant s$,

$$\left|\frac{\partial \theta_r}{\partial \zeta_i^r}(\zeta_r')\right| \leqslant L$$

成立.

其次, 由第 1.17 节中的定义通过简单的运算可以验证, 对于每个曲面 $\Theta(\omega_r) \subset \mathbb{R}^n$, 其中对每个 $\zeta_r' \in \omega_r, \Theta(\zeta_r') := (\zeta_r', \theta_r(\zeta_r'))$, 在这个特殊情况下其 $n-1$ 维面积由 $\int_{\omega_r} \left(1 + \sum_{i=1}^{n-1} \left|\frac{\partial \theta_r}{\partial \zeta_i^r}\right|^2\right)^{\frac{1}{2}} d\zeta_r'$ 给出, 这就验证了用于定义 $\int_\Gamma f d\Gamma$ 的表达式.

最后, 可以证明这样定义的数 $\int_\Gamma f d\Gamma$ 独立于局部坐标系的选取, 也与单位分解的取法无关.

Γ 的 $d\Gamma$ 可测子集 Γ_0 的面积定义并记作

$$d\Gamma\text{-meas}\, \Gamma_0 := \int_\Gamma \chi_{\Gamma_0} d\Gamma,$$

或者当 $n = 3$ 时

$$\text{area}\, \Gamma_0 = \int_\Gamma \chi_{\Gamma_0} d\Gamma.$$

这里 $\chi_{\Gamma_0} : \Gamma \to \mathbb{R}$ 表示集合 Γ_0 的特征函数.

函数 θ_r 几乎处处可微性的另一个重要推论是沿 $\Gamma, d\Gamma$ 几乎处处存在着单位外法向量场 $\boldsymbol{\nu} = (\nu_i)_{i=1}^n$. 这里的 “单位” “外” 分别表示对 $d\Gamma$ 几乎所有的 $x \in \Gamma$, $|\boldsymbol{\nu}(x)| = 1$, 且存在某 $\varepsilon(x) > 0$, $\{x + t\boldsymbol{\nu}(x); 0 \leqslant t < \varepsilon(x)\} \cap \Omega = \varnothing$, “法向” 表示 $\boldsymbol{\nu}(x)$ 正交于 Γ 的切超平面, 基于同样的理由, 后者也是 $d\Gamma$ 几乎处处存在的.

域的一个重要性质是以下基本 Green 公式成立, 它其实就是熟知的分部积分公式 $\int_a^b f'(t)g(t)dt = -\int_a^b f(t)g'(t)dt + f(b)g(b) - f(a)g(a)$ 的高维推广.

定理 1.18-2 (基本 Green 公式)　设 Ω 为 $\mathbb{R}^n$ 中的一个域, $\boldsymbol{\nu} = (\nu_i)_{i=1}^n$ 为沿 Ω 的边界 Γ 的单位外法向量场. 则对任何函数 $f, g \in C^1(\overline{\Omega})$,

$$\int_\Omega (\partial_i f) g dx = -\int_\Omega f \partial_i g dx + \int_\Gamma f g \nu_i d\Gamma,$$

其中 $1 \leqslant i \leqslant n$.

利用基本 Green 公式可以证明其他的 Green 公式, 其本质即是将 Ω 上积分的特殊组合表示为 Γ 上曲面积分的组合. 例如, 假设给定向量场 $\boldsymbol{v} = (v_i) \in \mathcal{C}^1(\overline{\Omega} \cdot \mathbb{R}^n)$, 则由基本 Green 公式, 对每个 $1 \leqslant i \leqslant n$, $\int_\Omega \partial_i v_i dx = \int_\Gamma v_i \nu_i d\Gamma$. 从而

$$\int_\Omega \operatorname{div} \boldsymbol{v} dx = \int_\Gamma \boldsymbol{v} \cdot \boldsymbol{\nu} d\Gamma,$$

其中 $\operatorname{div} \boldsymbol{v} := \sum_{i=1}^n \partial_i v_i$. 这个 Green 公式即向量场的散度定理.

第 2 章 赋范向量空间

引言

线性泛函分析构成了第 2 章到第 5 章讨论的主题.

具体而言, 本章的目标是给出完备或非完备的赋范向量空间的基本性质. 其后, 第 3 章用于讨论完备的赋范向量空间, 第 4 章讨论范数由内积导出的完备或非完备的赋范向量空间, 最后第 5 章将集中建立起这些空间上更为精巧而深刻的冠以“基本定理”的一些性质.

在本章介绍的主要概念中, 有关于连续线性或多重线性算子的 (2.9 和 2.11 节), 也有关于紧线性算子的 (2.10 节). 另一个关键的概念是紧性, 如漂亮的 F. Riesz 定理 (定理 2.7-3) 所展示的, 它给出了有限维的重要特征, 紧性在代数学基本定理 (定理 2.8-1) 的证明中起核心作用.

本章还将介绍无限维赋范向量空间的一些基本例子. 例如, 由紧集 K 到赋范向量空间 Y 的所有连续函数组成的空间 $\mathcal{C}(K;Y)$ (2.3 节), 当 $1 \leqslant p \leqslant \infty$ 时的空间 ℓ^p (2.4 节), 相应于 $1 \leqslant p \leqslant \infty$ 和 $\mathbb{R}^n$ 中任意开子集 Ω 的空间 $L^p(\Omega)$ (2.5 节), 以及由赋范向量空间 X 到赋范向量空间 Y 的所有连续线性算子组成的空间 $\mathcal{L}(X;Y)$ (2.9 节). 特别地, 在空间 $L^p(\Omega), 1 \leqslant p < \infty$, 中还利用光滑化算子的方法给出了由光滑函数作函数逼近的详尽处理.

作为应用部分, 本章包括了逼近论的一些基本结果. 例如, 由一般多项式 (定理 2.13-3 和 2.15-2) 或三角多项式 (定理 2.14-3) 逼近连续函数的 Weierstrass 逼近定理: 将 Korovkin 定理 (定理 2.12-1) 应用于 Bernstein 多项式 (定理 2.13-2) 或 Fejér 三角多项式 (定理 2.14-2), 给出了这些定理的构造性证明. 由这些结果还导出了更为一般且更为抽象的 Stone-Weierstrass 定理 (定理 2.15-1 和 2.15-3).

本章还包括关于凸性的介绍 (2.16 和 2.17 节). 这个概念在投影定理 (第 4 章)、Banach-Saks-Mazur 定理 (第 5 章)、极小化特征 (第 7 章) 和变分运算 (第 9 章) 中起关键的作用.

2.1 向量空间; Hamel 基; 向量空间的维数

在下文中, 用 $\mathbb{K}$ 表示数域 $\mathbb{R}$ 或数域 $\mathbb{C}$, 称 $\mathbb{K}$ 的元素为数量. 设 X 为一个集合, 如果存在两个映射:

$$(x,y) \in X \times X \to (x+y) \in X \quad \text{和} \quad (\alpha,x) \in \mathbb{K} \times X \to \alpha x \in X,$$

分别称之为*加法*和*数乘*, 满足如下性质: 对任何 $x,y,z \in X$,

$$x+y=y+x, \quad x+(y+z)=(x+y)+z.$$

存在 X 中的一个元, 记作 0, 使得对任何 $x \in X$, 有 $x+0=x$; 给定任何 $x \in X$, 存在 X 中一个元, 记作 $(-x)$, 使得 $x+(-x)=0$ (这样, 带有加法的集合 X 为一个 Abel 群); 对任何 $\alpha,\beta \in \mathbb{K}$ 和 $x,y \in X$,

$$\alpha(x+y)=\alpha x+\alpha y, \quad (\alpha+\beta)x=\alpha x+\beta x,$$
$$\alpha(\beta x)=(\alpha\beta)x, \quad 1x=x,$$

则称 X 称为 *$\mathbb{K}$ 上的向量空间*. 由上述的性质直接可导出下列结论: 0 是唯一的; 对任意给定的 $x \in X$, $(-x)$ 是唯一的; 对任意的 $x,y \in X, -(-x)=x, -(x+y)=(-x)+(-y)$; 对任意的 $\lambda \in \mathbb{K}$ 和 $x \in X, \lambda 0=0, 0x=0, (-\lambda)x=-(\lambda x)$; 如果 $x \neq 0$, 则当 $\lambda x=0$ 时 $\lambda=0$; 因为 $0 \in X$, 所以向量空间总是非空的; 因为加法满足结合律, 所以可记 $x+y+z := x+(y+z)$. 也常用简略的记号 $-x := (-x)$ 和 $x-y := x+(-y)$.

*实向量空间*即 $\mathbb{K}=\mathbb{R}$ 上的向量空间. *复向量空间*即 $\mathbb{K}=\mathbb{C}$ 上的向量空间. 向量空间或是实向量空间或为复向量空间.

集合 X 和 $\mathbb{K}$ 的元分别称作向量和数量. 元素 $0 \in X$ 称作 X 的*原点*或*零向量*; 注意, 同一个记号 0 既表示 X 中的零向量, 也表示 $\mathbb{K}$ 中的数量. 如果 $X \neq \{0\}$, 则任何 $x \in X$, 当满足 $x \neq 0$ 时, 称作 X 中的*非零向量*.

所谓 $\mathbb{K}$ 上向量空间 X 的*子空间*, 是指 X 的一个子集, 其本身也是 $\mathbb{K}$ 上的向量空间. 特别地, $\{0\}$ 是 X 的子空间. 如果 Y 是 X 的子空间, 且满足 $Y \neq X$, 则称 Y 是 X 的*真子空间*.

设 Y 和 Z 是向量空间 X 的两个子空间. 如果对每个 $x \in X$, 均可唯一地表示为其中 $y \in Y, z \in Z$,

$$x=y+z$$

则称 X 为 Y 和 Z 的*直接和*.

子空间的另一个例子是由 X 的子集 A *张成的子空间*, 它由 A 中向量的所有*有限线性组合*构成, 即由所有形如 $x=\sum_{j \in J}\alpha_j a_j$ 的元 $x \in X$ 构成, 其中指标集 J 是有限

的, 对所有 $j \in J, \alpha_j \in \mathbb{K}, a_j \in A$. 这个子空间记作

$$\mathrm{Span}\, A.$$

如果 X 的子集 A 形如 $A = \cup_{i=1}^n \{x_i\}$ 或 $A = \cup_{i\in I}\{x_i\}$, 则子空间 Span A 也记作

$$\mathrm{Span}\,(x_i)_{i=1}^n \quad 或 \quad \mathrm{Span}\,(x_i)_{i\in I}.$$

下面的概念是由 G. Hamel[1] 引进的 (其目的是为了解一个特殊的函数方程, 见习题 2.1-1). 设 $X \neq \{0\}$ 为向量空间, 所谓 X 的一个 Hamel 基是指任何一个由向量 $e_i \in X$ 组成的向量族 $(e_i)_{i\in I}$ (1.3 节), 满足以下两个性质:

第一, 这个向量族的元是线性无关的, 即对族 $(e_i)_{i\in I}$ 的任何有限子集 $(e_j)_{j\in J}$ 和任何一组数 $\alpha_j \in \mathbb{K}, j \in J$, 当 $\sum_{j\in J}\alpha_j e_j = 0$ 时, 有 $\alpha_j = 0, j \in J$. 第二, $\mathrm{Span}\,(e_i)_{i\in I} = X$, 即对任意给定的向量 $x \in X$, 存在族 $(e_i)_{i\in I}$ 的有限子族 $(e_j)_{j\in J(x)}$ 和数 $x_j \in \mathbb{K}, j \in J(x)$, 使得 $x = \sum_{j\in J(x)} x_j e_j$.

注意, 由第一个性质可知 Hamel 基中的所有向量 e_i, $i \in I$, 均是非零的, 且互不相等; 对给定的任何 $x \in X$, 相应的数 $x_j, j \in J(x)$ 是唯一确定的.

例如, 设 $e_n(x) = x^n, x \in \mathbb{R}$, 族 $(e_n)_{n=0}^\infty$ 组成了一元实变量多项式空间的 Hamel 基.

作为选择公理 (这里取 Zorn 引理的形式) 的第一个应用, 我们现在来证明在任何向量空间中, Hamel 基均是存在的, 其基数也是确定的. 另一个相关的性质系有限维空间中一条熟知的性质推广到一般向量空间的情况, 见习题 2.1-2.

定理 2.1-1 设 $X \neq \{0\}$ 为一个向量空间, 则

(a) 存在 X 的 Hamel 基.

(b) 设 E 和 F 为 X 的两个 Hamel 基, 则 $\mathrm{card}\, E = \mathrm{card}\, F$.

证明 (i) 用 $\mathcal{F}$ 表示 X 中向量的所有线性无关族组成的集合. 因为 $\mathcal{F}$ 包含 $\{e\}$, 其中 e 是 X 中任何非零向量, 所以 $\mathcal{F}$ 是非空的. 当 $\cup_{i\in I}\{e_i\} \subset \cup_{j\in J}\{e_j\}$ 时规定 $E = (e_i)_{i\in I} \preccurlyeq F = (e_j)_{j\in J}$, $\mathcal{F}$ 按关系 $\preccurlyeq$ 为偏序集. 因为族 $E = (e_i)_{i\in I}$ 恒同于子集 $\cup_{i\in I}\{e_i\}$ (线性无关族的元素互不相同), 关系 $E \preccurlyeq F$ 即简单的包含关系 $E \subset F$.

设 $\mathcal{E}$ 为 $\mathcal{F}$ 的全序子集. 于是族 $G := \cup_{E\in\mathcal{E}} E$ 是 $\mathcal{F}$ 中的元, 这是因为集合 $\mathcal{E}$ 是全序的, 故而 G 的任何有限子集 $(e_i)_{i=1}^m$ 是某个族 $E \in \mathcal{E}$ 的子族. 由此, 向量 e_i, $1 \leqslant i \leqslant m$, 是线性无关的. 由 G 的构造, 对任何 $E \in \mathcal{E}$, 有 $E \subset G$, 所以 G 显然是 $\mathcal{E}$ 的一个上界.

由 Zorn 引理 (定理 1.3-1), 集合 $\mathcal{F}$ 具有极大元 M, 它就是 X 的 Hamel 基. 因为否则必定存在非零向量 $e \in X$, e 不能表示为 M 中元的线性组合. 于是, $M \cup \{e\}$

[1] G. Hamel [1905]: Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x+y) = f(x) + f(y)$. Mathematische Annalen 60, 459–462.

是 $\mathcal{F}$ 中的元 (显然, $M \cup \{e\}$ 是线性无关族), 且满足 $M \prec M \cup \{e\}$, 此矛盾于 M 的极大性. 这就证明了 (a).

(ii) 再设 $E = \cup_{i \in I}\{e_i\}$ 和 $F = \cup_{j \in J}\{f_j\}$ 分别表示 X 的两个 Hamel 基. 特别地, 基 E 中每个元 e_i 可以表示为元 $f_j, j \in J(i)$ 的有限线性组合, 其中 $J(i)$ 是集合 J 的有限子集.

今断言 $F = \cup_{i \in I} F_i$, 其中 $F_i := \cup_{j \in J(i)}\{f_j\}$. 为此, 设存在 $j_0 \in J$, 使得 $f_{j_0} \notin \cup_{i \in I} F_i$. 因为 E 是基, 所以 $F - \{f_{j_0}\}$ 也是基, 此矛盾于 F 为基的假设, 所以 $F = \cup_{i \in I} F_i$.

先设两个基之一, 如 E, 为有限集. 由关系式 $F = \cup_{i \in I} F_j$ 可知基 F 也是有限集 (集合 I 和 $J(i)$, $i \in I$, 均为有限集). 此时, 有 $\text{card } E = \text{card } F$[2].

再设基 E, 或等价地集合 I, 是无限集. 则对每个 $i \in I$, 存在满射 $f_i; \mathbb{N} \to F_i$ (因集合 F_i 是有限集), 因而映射 $(i, n) \in I \times \mathbb{N} \to f_i(n) \in \cup_{i \in I} F_i = F$ 也是一个满射. 由此, $\text{card } F \preccurlyeq \text{card } (I \times \mathbb{N})$(定理 1.5-1). 因为 I 是无限集, $\text{card } \mathbb{N} \preccurlyeq \text{card } I$ (定理 1.5-3(a)), 所以

$$\text{card } (I \times \mathbb{N}) \preccurlyeq \text{card } (I \times I) = \text{card } I$$

(定理 1.5-3(b)). 因此, $\text{card } F \preccurlyeq \text{card } I = \text{card } E$. 同样可得 $\text{card } E \preccurlyeq \text{card } F$. 从而 $\text{card } E = \text{card } F$. □

如果向量空间 X 存在有限的或无限的 Hamel 基, 则称 X 为有限维或相应的无限维空间, 其维数即任意一个 Hamel 基的基数, 记作

$$\dim X$$

(由定理 2.1-1(b), 给定向量空间的任何两个 Hamel 基的基数相同, 因而这个定义是合理的). 有限维向量空间 X 的 Hamel 基简称为基.

Hamel 基将有限维向量空间的基的概念推广到任意的向量空间.

任意次 $n \geqslant 0$ 的实多项式 $p : x \in \mathbb{R} \to p(x) = \sum_{j=0}^{n} c_j x^j$ 组成的空间 $\mathcal{P}$ 就是一个无限维向量空间的例子, 这是因为族 $\mathcal{H} := (e_j)_{j=0}^{\infty}$ 显然是 $\mathcal{P}$ 的一个 Hamel 基, 其中 e_j 表示多项式 $x \in \mathbb{R} \to x^j, j \geqslant 0$, 称之为 $\mathcal{P}$ 的典则基. 此时 $\dim \mathcal{P} = \text{card } \mathcal{H} = \text{card } \mathbb{N}$.

注 不同的是任何无限维完备赋范向量空间的 Hamel 基 H 的基数总满足 $\text{card } H \succ \text{card } \mathbb{N}$ (定理 5.1-4).

习题

2.1-1 (1) 设集合 $\mathcal{F}$ 为所有对任何 $x, y \in \mathbb{R}$ 满足 $f(x+y) = f(x) + f(y)$ 的函数 $f : \mathbb{R} \to \mathbb{R}$ 的全体, 求 $\mathcal{F}$.

[2] 假定读者熟悉有限维空间的这一类性质.

提示: 视作域 $\mathbb{Q}$ 上的一个向量空间, 利用 $\mathbb{R}$ 的一个 Hamel 基.

(2) 集合 $\mathcal{F}$ 的基数是多少?

2.1-2 设 $X \neq \{0\}$ 为向量空间, $(e_j)_{j\in J}$ 为 X 中任何一族线性无关的向量. 证明存在 X 的 Hamel 基, 它以 $(e_j)_{j\in J}$ 为其一个子族.

2.2 赋范向量空间; 基本性质和例; 商空间

设 X 为 $\mathbb{K}$ 上的向量空间, 其中 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$. 设映射 $\|\cdot\| : X \to \mathbb{R}$ 满足下列性质:

$$\begin{aligned}
&\text{对任何 } x \in X, \|x\| \geqslant 0, \text{ 而 } \|x\| = 0 \text{ 当且仅当 } x = 0,\\
&\text{对任何 } \alpha \in \mathbb{K} \text{ 和 } x \in X,\ \|\alpha x\| = |\alpha|\,\|x\|,\\
&\text{对任何 } x, y \in X,\ \|x+y\| \leqslant \|x\| + \|y\|,
\end{aligned}$$

则称 $\|\cdot\|$ 为 X 上的范数. 上述最后一个不等式称为三角不等式. 称偶对 $(X, \|\cdot\|)$ 为赋范向量空间, 其中 X 为向量空间, $\|\cdot\|$ 为 X 上的范数.

有时, 我们也需要以下稍弱一些的概念. 设 X 为 $\mathbb{K}$ 上的向量空间, 映射 $|\cdot| : X \to \mathbb{R}$ 满足以下性质:

$$\begin{aligned}
&\text{对任何 } x \in X, |x| \geqslant 0,\\
&\text{对任何 } \alpha \in \mathbb{K} \text{ 和 } x \in X, |\alpha x| = |\alpha||x|,\\
&\text{对任何 } x, y \in X, |x+y| \leqslant |x| + |y|,
\end{aligned}$$

则称 $|\cdot|$ 为 X 上的半范数.

设 X 为赋范向量空间, 则有不等式: 对任何 $x, y \in X$,

$$\big|\,\|x\| - \|y\|\,\big| \leqslant \|x - y\|,$$

对任何 $x_i \in X,\ 1 \leqslant i \leqslant n$,

$$\left\|\sum_{i=1}^{n} x_i\right\| \leqslant \sum_{i=1}^{n} \|x_i\|.$$

由范数的定义立即可得:

定理 2.2-1 设 $(x, \|\cdot\|)$ 为赋范向量空间, 映射 $d : X \times X \to \mathbb{R}$ 定义为对任何 $x, y \in X$, $d(x, y) = \|x - y\|$, 则 d 是 X 的距离.

在赋范向量空间 $(X, \|\cdot\|)$ 上, 按上述距离 $d, (X, d)$ 为距离空间. X 上由此距离导出的拓扑 (1.10 节) 称为由范数 $\|\cdot\|$ 导出的拓扑, 或 X 上的范数拓扑, 也称作强拓扑.

除非有特别说明, 赋范向量空间总被视为赋以范数拓扑.

注 后面 (5.12 节) 将会看到无限维赋范向量空间还可赋以同样重要又与之不同的另一类拓扑, 即*弱拓扑*.

设 X 是赋以拓扑的向量空间. 如果这个拓扑可由 X 上的一个范数导出, 则称它是*可赋范的*. 习题 2.3-2 和 2.3-3 提供了一个向量空间上不可赋范的拓扑的例子.

下面定理中的范数是有限维向量空间上最常使用的, 它们提供了赋范向量空间的第一个例子. 习题 2.4-1 将对记号 $\|\cdot\|_\infty$ 再作讨论.

定理 2.2-2 *设 X 为 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$ 上的有限维向量空间, $(e_i)_{i=1}^n$ 为 X 的一个基.*

(a) *对每个扩充实数 $1 \leqslant p \leqslant \infty$, 定义映射 $\|\cdot\|_p$ 为*

$$x=\sum_{i=1}^n x_i e_i \in X \to \|x\|_p=\left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}, \quad 1 \leqslant p<\infty,$$

$$x=\sum_{i=1}^n x_i e_i \in X \to \|x\|_\infty=\max_{1 \leqslant i \leqslant n}|x_i|, \quad p=\infty,$$

则 $\|\cdot\|_p$ 为 X 上的范数.

(b) *对每个 $1 \leqslant p \leqslant \infty$, 空间 $(X,\|\cdot\|_p)$ 是可分的.*

证明 在 (a) 的证明中, 唯一不平凡的性质是当 $1<p<\infty$ 时的三角不等式, 它是后面定理 2.4-1 的证明中要建立的更一般的*关于序列的 Minkowski 不等式*的特殊情况.

为了对所有的 $1 \leqslant p \leqslant \infty$ 证明 (b), 只要注意到当 $\mathbb{K}=\mathbb{R}$ 时的可数无限集 $\{\sum_{i=1}^n y_i e_i \in X; y_i \in \mathbb{Q}, 1 \leqslant i \leqslant n\}$ 和当 $\mathbb{K}=\mathbb{C}$ 时的可数无限集 $\{\sum_{i=1}^n y_i e_i \in X; \text{Re } y_i \in \mathbb{Q}$ 且 $\text{Im } y_i \in \mathbb{Q}$, 其中 $1 \leqslant i \leqslant n\}$ 均是 $(X,\|\cdot\|_p)$ 中的稠密集, 因此 $(X,\|\cdot\|_p)$ 是可分空间. □

注 后面我们将会看到, 实际上任何有限维赋范向量空间均是可分的 (定理 2.7-1).

注意, 与这些范数相应的距离 $d_p: X\times X \to \mathbb{R}, 1 \leqslant p \leqslant \infty$, 即对任何 $(x,y) \in X\times X$, 由 $d_p(x,y)=\|x-y\|_p$ 定义的距离, 正是在第 1.10 节中引进的距离, 称范数 $\|\cdot\|_2$ 为 *Euclid 范数*, 在不致发生混淆的情况下, 简记

$$|\cdot|:=\|\cdot\|_2.$$

对每个整数 $n \geqslant 2$, 在由数量 $x_i \in \mathbb{K}$ 的 n 元组 $(x_i)_{i=1}^n$ 的全体组成的向量空间 $\mathbb{K}^n$ 上, 赋以某一个范数 $\|\cdot\|_p, 1 \leqslant p \leqslant \infty$, 则它就成为一个赋范向量空间, 任何一个这类范数在 $\mathbb{K}^n$ 上导出的拓扑即 $\mathbb{K}^n$ 上的通常拓扑 (1.10 节), 因而后者是可赋范的 (可以在由 $n\times n$ 阵组成的向量空间上定义类似的拓扑, 因为它等同于空间 $\mathbb{K}^{n^2}$; 见习题 2.2-1).

另一个赋范向量空间的例子是同一个域 $\mathbb{K}$ 上的赋范向量空间的乘积空间 $X = X_1 \times X_2 \times \cdots \times X_n$, X 上可赋予下列范数之一

$$x = (x_j)_{j=1}^n \to \left(\sum_{j=1}^n \|x_j\|_{X_j}^p\right)^{\frac{1}{p}}, \text{其中 } 1 \leqslant p < \infty,$$

$$x = (x_j)_{j=1}^n \to \max_{1 \leqslant j \leqslant n} \|x_j\|_{X_j},$$

这里每一个范数导出 X 上的乘积拓扑.

设 X 为 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的向量空间, Z 为 X 的子空间. 容易验证, 关系

$$x \sim y \quad \text{当且仅当 } (x - y) \in Z$$

是 X 上的一个等价关系 (1.1 节). 用

$$\begin{aligned}[x] &= \{y \in X; (x - y) \in Z\} \\ &= \{(x - z) \in X; z \in Z\} \subset \mathcal{P}(X)\end{aligned}$$

记 x 模这个关系的等价类.

容易知道, 如果在商集 X/Z(即由上述等价类组成的集合, 见 1.1 节) 上定义加法和数乘: 对一切 $x, y \in X$ 和 $\alpha \in \mathbb{K}$,

$$[x] + [y] = [x + y], \quad \alpha[x] = [\alpha x],$$

则 X/Z 也成为 $\mathbb{K}$ 上的一个向量空间, 称之为商空间, X/Z 中的零向量即 $[0] = Z$. 当对空间 Z 的定义无歧义时, 也使用记号

$$[X] := X/Z.$$

注 在其他一些地方, $x \in X$ 的等价类 $[x]$ 和商空间 $[X]$ 也被记作 $\dot{x}$ 和 $\dot{X}$.

例如, 设 e_1, e_2, e_3 是空间 $\mathbb{R}^3$ 的典范基. 则商空间 $\mathbb{R}^3/\text{Span } e_1$ 是由平行直线 Span e_1 的所有直线组成的 (实) 向量空间; 商空间 $\mathbb{R}^3/\text{Span } (e_1, e_2)$ 是由平行于平面 Span (e_1, e_2) 的所有平面组成的 (实) 向量空间.

如果 X 是赋范向量空间, Z 是 X 的闭子空间, 则商空间 X/Z 提供了另一个赋范向量空间的例子.

定理 2.2-3 设 $(X, \|\cdot\|_X)$ 为赋范向量空间, Z 为 X 的闭子空间, 定义映射 $\|\cdot\| : X/Z \to \mathbb{R}$ 为

$$\|[x]\| := \inf_{y \in [x]} \|y\|_X = \inf_{z \in Z} \|x - z\|_X,$$

则 $\|\cdot\|$ 是商空间 X/Z 上的范数, 称为商范数.

证明 显然地, 对任何 $[x] \in X/Z, \|[x]\| \geqslant 0$, 且 $\|[0]\| = 0$. 如果 $[x] \in X/Z$ 使得 $\|[x]\| = \inf_{z\in Z} \|x - z\|_X = 0$, 则 $x \in \overline{Z}$, 而 Z 是闭的, 故 $x \in Z$, 这样, $[x] = Z$, 即它就是 X/Z 中的零向量. 另外, 对任何 $\alpha \in \mathbb{K}, x, y \in X$,

$$\begin{aligned}
\|\alpha[x]\| = \|[\alpha x]\| &= \inf_{z\in Z} \|\alpha x - z\|_X \\
&= \inf_{u\in Z} \|\alpha(x-u)\|_X = |\alpha| \inf_{z\in Z} \|x - z\|_X = |\alpha|\, \|[x]\|, \\
\|[x] + [y]\| = \|[x+y]\| &= \inf_{z\in Z} \|x + y - z\|_X \\
&= \inf_{u,v\in Z} \|(x-u) + (y-v)\|_X \\
&\leqslant \inf_{u,v\in Z} (\|x-u\|_X + \|y-v\|_X) \\
&= \inf_{u\in Z} \|x-u\|_X + \inf_{v\in Z} \|y-v\|_X = \|[x]\| + \|[y]\|.
\end{aligned}$$

这就证得了 $\|\cdot\|$ 是商空间的范数. □

因为赋范向量空间 $(X, \|\cdot\|)$ 既是拓扑空间, 也是距离空间, 所以它继承了第 1 章中所复述的距离空间的所有定义和性质. 特别地, 对于某个 $x \in X$ 和 $r > 0$, 以 x 为中心, r 为半径的球是 X 中形如

$$B(x; r) = \{y \in X; \|y - x\| < \gamma\}$$

的子集. 它是 X 中的开集 (定理 1.10-1). 单位球是指特殊的球

$$B(0; 1) = \{x \in X; \|x\| < 1\}.$$

X 中的子集 A 是开集当且仅当对任意给定的点 $x \in A$, 存在包含于 A 中的球 $B(x; \gamma)$.

X 中向量 x_n 的序列 $(x_n)_{n=0}^{\infty}$ 收敛于 $x \in X$ 是指当 $n \to \infty$ 时

$$\|x_n - x\| \to 0.$$

顺便注意, 特别当这里的 "范数收敛" 须区分于弱收敛 (5.12 节) 时, (如上定义的) 收敛序列也称为强收敛.

设 A 是 $(X, \|\cdot\|)$ 的子集, 如果存在数 M, 对所有的 $x \in A$, 均有 $\|x\| \leqslant M$, 则称 A 为有界的.

相对而言, 下面的概念是赋范向量空间特有的. 给定任何 $x \in X$ 和 $r > 0$, (开) 球 $B(x, r)$ 的闭包

$$\overline{B(x, r)} = \{y \in X; \|y - x\| \leqslant r\}$$

称为以 x 为中心, r 为半径的闭球, 当 $x = 0, r = 1$ 时, 简称为闭单位球, 球 $B(x, r)$ 的边界

$$\partial B(x, r) = \{y \in X; \|y - x\| = r\}$$

称为以 x 为中心, r 为半径的球面, 当 $x=0, r=1$ 时, 简称为单位球面.

因为赋范向量空间 X 赋有两种特殊的运算, 其距离 (如定理 2.2-1 所构造) "关于这些运算是相容的", 即对任何 $x, y, z \in X$ 和任何 $\lambda \in \mathbb{K}$, $d(x+z, y+z) = d(x,y)$, $d(\lambda x, \lambda y) = |\lambda| d(x,y)$, 因而空间 X 比任意的距离空间, 当然也比任意的拓扑空间, 具有更多的性质. 相应地, 在本章和下一章中, 我们将研究对于赋范向量空间而言更为特殊的拓扑或距离及其他相关的性质.

关于这一方面, 我们从一个定义开始: 在给定的向量空间 X 上, 两个范数 $\|\cdot\|$ 和 $\|\cdot\|'$ 称为等价的是指在 X 上由 $\|\cdot\|$ 和 $\|\cdot\|'$ 导出的拓扑相同, 下面的定理给出了两个范数等价性的简单而又基本的准则.

定理 2.2-4 向量空间 X 上两个范数 $\|\cdot\|$ 和 $\|\cdot\|'$ 等价, 当且仅当存在常数 C 和 C', 使得对任何 $x \in X$,

$$\|x\|' \leqslant C\|x\|, \quad \|x\| \leqslant C'\|x\|'.$$

证明 (i) 假定 $\|\cdot\|$ 和 $\|\cdot\|'$ 为等价范数. 则恒等映射 $\mathrm{id}: (X, \|\cdot\|) \to (X, \|\cdot\|')$ 是连续的 (因为开集相同, 见定理 1.7-3). 特别地, $(X, \|\cdot\|')$ 中开集 $B' := \{y \in X; \|y\|' < 1\}$ 的逆像 $\mathrm{id}^{-1}(B')$ 是 $(X, \|\cdot\|)$ 中包含 0 的开集 (因为 $I(0) = 0 \in B'$). 因此, 存在常数 $C > 0$, 使得集合 $\{y \in X; \|y\| < \frac{1}{C}\}$ 的闭包包含于 $\mathrm{id}^{-1}(B')$ 中. 从而

$$\text{当 } \|y\| \leqslant \frac{1}{C} \text{ 时, } \|y\|' \leqslant 1.$$

给定任何非零向量 $x \in X$, 向量 $y := \frac{1}{C\|x\|}x$ 满足要求 $\|y\| = \frac{1}{C}$, 因此 $\|y\|' = \frac{1}{C\|x\|}\|x\|' \leqslant 1$. 于是, 对所有的 $x \in X$. 不等式 $\|x\|' \leqslant C\|x\|$ 成立. 同样可得另一个不等式.

(ii) 假定对所有的 $x \in X, \|x\|' \leqslant C\|x\|$. 由此不等式可知距离空间 $(X, \|\cdot\|')$ 中任何中心在 $y \in X$, 半径为 r 的球的闭包包含一个距离空间 $(X, \|\cdot\|)$ 中的中心在 $y \in X$, 半径为 $\frac{r}{C}$ 的球. 因而, 由 $\|\cdot\|'$ 导出的拓扑下的任何开集在由 $\|\cdot\|$ 导出的拓扑下是开集. 同样可得另一蕴涵关系. □

下面的定理列出了赋范向量空间 (赋以范数拓扑) 的另一些初等而常用的性质.

定理 2.2-5 设 X 为 $\mathbb{K}$ 上的赋范向量空间, 则映射

$$\|\cdot\|: x \in X \to \|x\| \in \mathbb{R},$$

$$(x,y) \in X \times X \to (x+y) \in X,$$

$$(\alpha, x) \in \mathbb{K} \times X \to \alpha x \in X$$

均是连续的.

证明 由不等式

$$|\,\|x\| - \|\tilde{x}\|\,| \leqslant \|x - \tilde{x}\|$$

即得映射 $x \in X \to \|x\| \in \mathbb{R}$ 的连续性. 后两个映射的连续性则由不等式

$$\|(x+y)-(\tilde{x}+\tilde{y})\| \leqslant \|x-\tilde{x}\|+\|y-\tilde{y}\|,$$
$$\|\alpha x-\tilde{\alpha}\tilde{x}\| \leqslant |\tilde{\alpha}|\,\|x-\tilde{x}\|+|\alpha-\tilde{\alpha}|\,\|\tilde{x}\|+|\alpha-\tilde{\alpha}|\,\|x-\tilde{x}\|$$

结合乘积拓扑的定义 (1.6 节) 及收敛序列的有界性 (对最后一个映射而言) 导出. □

所谓拓扑向量空间是指一个向量空间, 其上赋以使得加法和数乘均为连续映射的拓扑. 这样, 定理 2.2-5 证明了赋范向量空间是拓扑向量空间.

作为定理 2.2-5 的第一个应用, 我们来对于赋范向量空间的开子集, 建立一个有趣而特殊的性质 (这个性质对于任意的拓扑空间未必成立). 这里要用到的连通概念已在 1.9 节中给出.

定理 2.2-6 设 X 为赋范向量空间, A 为 X 中的开子集, 则 A 的连通分支是 X 中的开集.

证明 设 C 是 A 的一个连通分支, $x \in C$, 因为 $C \subset A$, A 是开集, 所以存在一个包含于 A 中的球 $B(x;r)$.
给定 $y, z \in B(x;r)$, 定义映射 $\gamma : [0,1] \to X$ 为

$$\gamma(\lambda) := (1-\lambda)y+\lambda z,\ 0 \leqslant \lambda \leqslant 1.$$

因为对任何 $\lambda \in [0,1]$,

$$\|\gamma(\lambda)-x\| = \|(1-\lambda)(y-x)+\lambda(z-x)\|$$
$$\leqslant (1-\lambda)\|y-x\|+\lambda\|z-x\| < r,$$

所以 γ 映区间 $[0,1]$ 到 $B(x;\gamma)$ 中. 由定理 2.2-5, 映射 $\gamma : [0,1] \to B(x;r)$ 是连续的.

因为 γ 是联结 y 和 z 的道路, 所以 $B(x;\gamma)$ 是道路连通的, 从而是连通的. 作为 A 中包含 x 的最大连通集, 集 C 必定包含 $B(x;\gamma)$. 因此, C 是开集. □

可分的赋范向量空间具有一个有趣的性质 (本书的下文中经常要用到):

定理 2.2-7 设 X 为可分的赋范向量空间. 则存在 X 的有限维子空间的可数无限族 $(X_n)_{n=1}^{\infty}$, 满足

$$\dim X_n = n, \quad X_n \subset X_{n+1}, \quad n \geqslant 1 \text{ 且 } \overline{\bigcup_{n=1}^{\infty} X_n} = X.$$

证明 设 $x_k \in X$, $k \geqslant 1$, 满足

$$\overline{\bigcup_{k=1}^{\infty}\{x_k\}} = X.$$

首先, 注意到不失一般性可设对任何 $k \geqslant 1, x_k \neq 0$ (否则, 取 $\widetilde{x_k} \in X, k \geqslant 1$, 满足对任何 $k \geqslant 1, \widetilde{x_k} \neq 0$, 且当 $k \to \infty$ 时, $\widetilde{x_k} \to 0$. 于是, 可数族 $(\cup_{k\in K}\{x_k\})\cup(\cup_{k=1}^{\infty}\{\widetilde{x_k}\})$ 在 X 中稠密). 此时, 递推地定义向量 $e_k \in X, k \geqslant 1$ 为

$$e_k := x_{\sigma(k)},$$

其中 $\sigma(1) := 1$, 当 $k \geqslant 2$ 时, $\sigma(k) := \min\{m \geqslant \sigma(k-1)+1; x_m \bar{\in} \text{Span }(e_l)_{l=1}^{k-1}\}$.

于是, 定义为

$$X_n := \text{Span }(e_k)_{k=1}^{n}$$

的子空间显然具有所要求的全部性质 ($\overline{\cup_{n=1}^{\infty}X_n} = X$ 由包含关系 $\cup_{k=1}^{\infty}\{x_k\} \subset \cup_{n=1}^{\infty}X_n$ 导出). □

在结束本节时, 我们来证明任何向量空间均是可赋范的. 这个结果的一般性自有其意义, 在它的证明中实际上借助于 (通过定理 2.1-1) 选择公理.

定理 2.2-8 任意向量空间均可赋范.

证明 给定任意一个 $\mathbb{K}$ 上向量空间 X, 设 $(e_i)_{i\in I}$ 为 X 的 Hamel 基 (定理 2.1-1). 对任意给定的向量 $x \in X$, 唯一地存在 I 的有限子集 $I(x)$ 和唯一确定的数 $x_i \in \mathbb{K}, i \in I(x)$, 使得 $x = \sum_{i\in I(x)} x_i e_i$. 可以直接验证映射

$$x = \sum_{i\in I(x)} x_i e_i \in X \to \sum_{i\in I(x)} |x_i|$$

是 X 上的一个范数. □

习题

2.2-1 (1) 证明所有 n 阶可逆实矩阵的集合在所有 n 阶实矩阵的集合 $\mathbb{M}^n$ 中是开的, 后者等同于赋以通常拓扑的空间 $\mathbb{R}^{n^2}$.

(2) 证明所有 n 阶实对称阵的集合 $\mathbb{S}^n$ 在 $\mathbb{M}^n$ 中是闭的.

(3) 证明所有 n 阶实对称且正定的矩阵的集合 $\mathbb{S}_>^n$ 在 $\mathbb{S}^n$ 中是开的, 其中后者赋以 $\mathbb{M}^n$ 的诱导拓扑.

(4) 作为 $\mathbb{M}^n$ 的子集, $\mathbb{S}_>^n$ 有什么性质?

(5) 证明 $\{A \in \mathbb{M}^n; \det A > 0\}$ 是 $\mathbb{M}^n$ 的连通子集.

2.2-2 证明赋范向量空间中的任何连通开子集均是道路连通的 (1.9 节).

2.2-3 判别下述命题成立与否: 设 $\|\cdot\|$ 和 $\|\cdot\|'$ 为同一个向量空间 X 上的两个范数, 若 $(x_n)_{n=1}^{\infty}$ 为诸元素 $x_n \in X$ 的一个序列, 且在 $(X, \|\cdot\|)$ 中 $\lim\limits_{n\to\infty} x_n = x$, 在 $(X, \|\cdot\|')$ 中 $\lim\limits_{n\to\infty} x_n = x'$, 则 $x = x'$.

2.2-4 设 K 为赋范向量空间 $(X, \|\cdot\|)$ 中的一个紧子集.

(1) 证明: 对任意给定的 $x \in X$, 存在 $y \in K$, 使得 $\|x-y\| = \inf\limits_{z\in K} \|x-z\|$.

(2) 如果对每个 $x \in X$, 上述 y 均是唯一的, 定义映射 $P: X \to K$, 使得对任何 $x \in X$, $\|x - Px\| = \inf_{z \in K} \|x - z\|$, 证明 P 是连续的.

2.3 K 为紧集时的空间 $\mathcal{C}(K;Y)$; 一致收敛和局部一致收敛性

本节来构造赋范向量空间另一个基本的例子, 即紧集上连续函数组成的空间, 其他基本的例子, 如空间 ℓ^p (2.4 节), $L^p(\Omega)$ (2.5 节), $1 \leqslant p \leqslant \infty$ 或商空间, 将在今后陆续给出.

如 $\mathcal{C}(K;Y)$ 或 $\mathcal{C}(K)$ 等记号已在 1.7 节中定义.

定理 2.3-1 设 K 为紧拓扑空间, $(Y, \|\cdot\|)$ 为赋范向量空间, 则 $\mathcal{C}(K;Y)$ 为向量空间, 函数 $|||\cdot||| : \mathcal{C}(K;Y) \to \mathbb{R}$ 定义为

$$|||f||| := \sup_{x \in K} \|f(x)\|, \quad f \in \mathcal{C}(K;Y),$$

$|||\cdot|||$ 为 $\mathcal{C}(K;Y)$ 上的范数.

证明 显然 $\mathcal{C}(K;Y)$ 为向量空间. 又因为 K 是紧的, 且作为连续函数的复合映射, 函数 $x \in K \to \|f(x)\|$ 是连续的 (定理 1.7-2 和 2.2-5), 由定理 1.13-6 易知 $\sup_{x \in K} \|f(x)\| < \infty$. 即可证明 $|||\cdot|||$ 为范数. □

由定理 2.3-1 引入的范数称为 sup 范数. 函数 $f_n \in \mathcal{C}(K;Y)$ 的序列 $(f_n)_{n=1}^{\infty}$ 称为 (当 $n \to \infty$ 时) 一致收敛于 $f \in \mathcal{C}(K;Y)$, 是指 $\lim_{n\to\infty} |||f_n - f||| = 0$, 即

$$\lim_{n \to \infty} (\sup_{x \in K} \|f_n(x) - f(x)\|) = 0.$$

在 $Y = \mathbb{R}$ 或 $Y = \mathbb{C}$ 这一重要的特例中, sup 范数简记为 $\|\cdot\|$. 即对所有的 $f \in \mathcal{C}(K)$ 或所有的 $f \in \mathcal{C}(K;\mathbb{C})$,

$$\|f\| := \sup_{x \in K} |f(x)|.$$

当 Ω 是 $\mathbb{R}^n$ 的有界开子集时, 记 $\|\cdot\|$ 为 sup 范数, 即定义为对所有的 $f \in \mathcal{C}(\overline{\Omega})$,

$$\|f\| := \sup_{x \in \overline{\Omega}} |f(x)|,$$

$(\mathcal{C}(\overline{\Omega}), \|\cdot\|)$ 提供了这类空间的基本例子, 其中包括 $n = 1$, $\Omega =]a, b[\subset \mathbb{R}$ 的情况 (在本章中将多次提及). 此时, 为简明起见, 记

$$\mathcal{C}[a, b] := \mathcal{C}([a, b]).$$

注 相应地, 当 Ω 是 $\mathbb{R}^n$ 中的任意开子集 (有界或无界时), "似乎类同的" 空间 $\mathcal{C}(\Omega)$ 实际上完全不同, 因为其上的 "自然" 拓扑尽管可距离化, 但却是不可赋范的 (参见习题 2.3-2).

类似地, 显然对每个整数 $m \geqslant 1$, 空间

$$\mathcal{C}^m(\overline{\Omega}) := \{f \in \mathcal{C}^m(\Omega),\ \text{对每个}\ |\boldsymbol{\alpha}| \leqslant m,\ \text{存在}\ g^{\boldsymbol{\alpha}} \in \mathcal{C}(\overline{\Omega}),\ \text{使得}\ \partial^{\boldsymbol{\alpha}} f = g^{\boldsymbol{\alpha}}|_{\Omega}\}$$

(1.18 节), 其中 Ω 是 $\mathbb{R}^n$ 中的有界开子集, 是赋范向量空间, 这里的范数 $\|\cdot\|_{\mathcal{C}^m(\overline{\Omega})}$ 定义为对任何 $f \in \mathcal{C}^m(\overline{\Omega})$,

$$\begin{aligned}\|f\|_{\mathcal{C}^m(\overline{\Omega})} :&= \max_{0\leqslant|\boldsymbol{\alpha}|\leqslant m} \sup_{x\in\overline{\Omega}} |g^{\boldsymbol{\alpha}}(x)| \\ &= \max_{0\leqslant|\boldsymbol{\alpha}|\leqslant m} \sup_{x\in\overline{\Omega}} |\partial^{\boldsymbol{\alpha}} f(x)|.\end{aligned}$$

实际上, 一致收敛的概念并非限于讨论定义于紧空间而取值于赋范向量空间的连续映射 (如定理 2.3-1 所涉). 更一般地, 设 X 为任一集合, Y 为赋范向量空间. 映射 $f_n : X \to Y$ 的序列 $(f_n)_{n=1}^{\infty}$ 称为当 $n \to \infty$ 时在 X 上一致收敛于映射 $f : X \to Y$, 是指

$$\lim_{n\to\infty} (\sup_{x\in X} \|f_n(x) - f(x)\|) = 0.$$

注意在这个定义中, 函数 $f_n, n \geqslant 1$ 和 f 可能是无界的. 例如, 考察函数 $f_n : x \in]0,\infty[\to \frac{1}{x} + \frac{1}{n}, n \geqslant 1$ 和 $f : x \in]0,\infty[\to \frac{1}{x}$ 即可.

如果函数 $f_n, n \geqslant 1$ 和 f 均为有界的, 则根据以下结果, 一致收敛可更为一般地视为按范数拓扑下的收敛 (易于直接证明, 此处从略).

定理 2.3-2 设 X 为任意集合, $(Y, \|\cdot\|)$ 为赋范向量空间, 则所有有界映射 $f : X \to Y$, 即其直接像 $f(X)$ 为 Y 中的有界子集的映射, 组成的集合

$$\mathcal{B}(X;Y)$$

为一个向量空间; 定义 $|||\cdot||| : \mathcal{B}(X;Y) \to \mathbb{R}$ 为对每个 $f \in \mathcal{B}(X;Y)$

$$|||f||| := \sup_{x\in X} \|f(x)\|,$$

则 $|||\cdot|||$ 是 $\mathcal{B}(X;Y)$ 上的范数.

这个概念还能进一步做如下推广: 设 X 为拓扑空间, Y 为赋范向量空间, 映射 $f_n : X \to Y$ 的序列 $(f_n)_{n=1}^{\infty}$ 称为当 $n \to \infty$ 时局部一致收敛于映射 $f : X \to Y$, 是指对任何 $x_0 \in X$, 存在 x_0 的邻域 $V(x_0)$, 使得

$$\lim_{n\to\infty} (\sup_{x\in V(x_0)} \|f_n(x) - f(x)\|) = 0.$$

这里的函数 $f_n, n \geqslant 1$, 和 f 也可能是无界的. 例如, 可考察函数 $f_n : x \in]0, \infty[\to \frac{1}{x} + \max\{0, x-n\}$; 则函数列 $(f_n)_{n=1}^{\infty}$ 局部一致 (非一致) 收敛于函数 $f : x \in]0, \infty[\to \frac{1}{x}$.

自然, 上述各种一致收敛的情况均可导致序列 $(f_n)_{n=1}^{\infty}$ 当 $n \to \infty$ 时点态收敛于 f, 即对任何 $x \in X$, 当 $n \to \infty$ 时

$$f_n(x) \to f(x).$$

局部一致收敛的关键性质在于它能保持连续性.

定理 2.3-3 *设 X 为拓扑空间, Y 为赋范向量空间, 当 $n \to \infty$ 时, 映射 $f_n : X \to Y$ 的序列 $(f_n)_{n=1}^{\infty}$ 局部一致收敛于映射 $f : X \to Y$. 如果映射 $f_n, n \geqslant 1$ 在点 $x_0 \in X$ 连续, 或在 X 上连续, 则 f 在 x_0 处连续, 或相应地在 X 上连续.*

证明 设映射 $f_n, n \geqslant 1$ 在点 $x_0 \in X$ 处连续. 又设 $\varepsilon > 0$. 由序列 $(f_n)_{n=1}^{\infty}$ 局部一致收敛, 从而存在 x_0 的邻域 $V(x_0)$, 使得 $\lim_{n\to\infty}(\sup_{x\in V(x_0)} \|f_n(x) - f(x)\|) = 0$. 于是, 可取 $n_0 \geqslant 1$, 使得

$$\sup_{x\in V(x_0)} \|f_{n_0}(x) - f(x)\| \leqslant \frac{\varepsilon}{3}.$$

映射 f_{n_0} 在 x_0 处连续, 从而存在 x_0 的邻域 $W(x_0) \subset V(x_0)$, 使得当 $x \in W(x_0)$ 时

$$\|f_{n_0}(x) - f_{n_0}(x_0)\| \leqslant \frac{\varepsilon}{3}.$$

于是, 对任何 $x \in W(x_0)$,

$$\|f(x) - f(x_0)\| \leqslant \|f(x) - f_{n_0}(x)\| + \|f_{n_0}(x) - f_{n_0}(x_0)\| + \|f_{n_0}(x_0) - f(x_0)\| \leqslant \varepsilon,$$

即映射 f 在 x_0 处连续.

由同样的讨论即知, 当 $f_n, n \geqslant 1$, 在 X 上所有的点处连续时, f 也在 X 上所有的点处连续. □

习题

2.3-1 (Dini 定理[3]) 给定紧距离空间 K, 函数 $f_n \in \mathcal{C}(K)$ 的单调递增序列 $(f_n)_{n=1}^{\infty}$ (当 $n \leqslant m$ 时, 对所有的 $x \in K, f_n(x) \leqslant f_m(x)$) 点点收敛于函数 $f \in \mathcal{C}(K)$. 证明 $(f_n)_{n=1}^{\infty}$ 一致收敛于 f.

2.3-2 以下设 Ω 是 $\mathbb{R}^n$ 的开子集, 给定任何函数 $f \in \mathcal{C}(\Omega)$ 和 Ω 的任何紧子集 K, 令

$$|f|_K := \sup_{x\in K} |f(x)|.$$

则易知如上定义的映射 $|\cdot|_K : \mathcal{C}(\Omega) \to \mathbb{R}$ 是空间 $\mathcal{C}(\Omega)$ 上的一个半范数, 但并非范数.

[3] U. Dini [1878]: Fondamenti per la Teoria delle Funzioni di Variabili Reali. T. Nistri, Pisa.

(1) 证明存在 Ω 的紧子集 K_i 的序列 $(K_i)_{i=1}^{\infty}$, 使得

$$K_i \subset \operatorname{int} K_{i+1}, \quad i \geqslant 1,$$
$$\Omega = \bigcup_{i=1}^{\infty} K_i.$$

(2) 设对任何 $i \geqslant 1, \alpha_i > 0$, 级数 $\sum_{i=1}^{\infty} \alpha_i$ 收敛. 对任何两个函数 $f, g \in \mathcal{C}(\Omega)$, 令

$$d(f,g) := \sum_{i=1}^{\infty} \alpha_i \frac{|f-g|_{K_i}}{1+|f-g|_{K_i}}.$$

证明如上定义的映射 $d : \mathcal{C}(\Omega) \times \mathcal{C}(\Omega) \to \mathbb{R}$ 是向量空间 $\mathcal{C}(\Omega)$ 上的距离.

(3) 证明: 在距离空间 $(\mathcal{C}(\Omega), d)$ 中, 函数 $f_n \in \mathcal{C}(\Omega)$ 的序列 $(f_n)_{n=1}^{\infty}$ 收敛于函数 $f \in \mathcal{C}(\Omega)$, 当且仅当对任何紧子集 $K \subset \Omega$,

$$\lim_{n\to\infty} |f_n - f|_K = 0.$$

(4) 距离空间 $(\mathcal{C}(\Omega), d)$ 是否完备?

(5) 证明空间 $\mathcal{C}(\Omega)$ 上由 (2) 给定的距离 d 导出的拓扑是不可赋范的.

空间 $\mathcal{C}(\Omega)$ 上由上述距离 d 导出的拓扑称为相应于半范数 $|\cdot|_K$ 族 $(|\cdot|_{K\in\mathcal{K}})$ 的 Fréchet 拓扑, 其中 $\mathcal{K}$ 表示 Ω 的所有紧子集族.

注 可以在 Ω 上 m 阶连续可微的函数空间上定义类似的 Fréchet 拓扑, 见习题 7.8-3.

2.3-3 对任意给定的函数 $f \in \mathcal{C}^{\infty}[0,1]$ 和任意整数 $n \geqslant 0$, 令

$$\|f\|_n := \max_{0\leqslant m\leqslant n} \sup_{0\leqslant x\leqslant 1} |f^{(m)}(x)|.$$

(1) 设对任何 $n \geqslant 1$ 均有 $\alpha_n > 0$, 且级数 $\sum\limits_{n=1}^{\infty} \alpha_n$ 收敛. 对任意两个函数 $f, g \in \mathcal{C}^{\infty}[0,1]$, 令

$$d(f,g) := \sum_{n=1}^{\infty} \alpha_n \frac{\|f-g\|_n}{1+\|f-g\|_n}.$$

证明上述定义的映射 $d : \mathcal{C}^{\infty}[0,1] \times \mathcal{C}^{\infty}[0,1] \to \mathbb{R}$ 是空间 $\mathcal{C}^{\infty}[0,1]$ 上的距离.

(2) 证明距离空间 $(\mathcal{C}^{\infty}[0,1], d)$ 是完备的.

(3) 证明在空间 $\mathcal{C}^{\infty}[0,1]$ 上由距离 d 导出的拓扑是不可赋范的.

2.3-4 设 X 为拓扑空间, $(Y, \|\cdot\|)$ 为赋范向量空间, 又设 $(\mathcal{B}(X;Y), |||\cdot|||)$ 为由定理 2.3-2 定义的赋范向量空间, 证明 $\mathcal{B}(X;Y) \cap \mathcal{C}(X;Y)$ 是 $(\mathcal{B}(X;Y), |||\cdot|||)$ 的闭子空间.

2.4 空间 ℓ^p, $1 \leqslant p \leqslant \infty$

从定理 2.2-2 可知, 由对任何 $x = (x_i)_{i=1}^{n} \in K^n$, 当 $1 \leqslant p < \infty$ 时令 $\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}}$, 当 $p = \infty$ 时令 $\|x\|_{\infty} = \max_{1\leqslant i\leqslant n} |x_i|$, 定义的映射 $\|\cdot\|_p$ 是 $\mathbb{K}^n$ 上的范数. 下面的定理提供了将上述范数推广到由 $x_i \in \mathbb{K}$ 的无穷序列 $(x_i)_{i=1}^{\infty}$ 组成的向量空间上的途径.

定理 2.4-1 (序列的 Hölder 不等式和 Minkowski 不等式) (a) 给定实数 $p>1$, 实数 q 满足 $\frac{1}{p}+\frac{1}{q}=1$ (因而 $q>1$), 又设数列 $x=(x_i)_{i=1}^{\infty}$ 和数列 $y=(y_i)_{i=1}^{\infty}$ 满足

$$\sum_{i=1}^{\infty}|x_i|^p<\infty,\quad \sum_{i=1}^{\infty}|y_i|^q<\infty,$$

则级数 $\sum_{i=1}^{\infty}|x_iy_i|$ 收敛, 而且以下的 Hölder 不等式[4)]成立:

$$\sum_{i=1}^{\infty}|x_iy_i|\leqslant\left(\sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{\infty}|y_i|^q\right)^{\frac{1}{q}}.$$

(b) 给定实数 $p\geqslant 1$, 又设数列 $x=(x_i)_{i=1}^{\infty}$ 和 $y=(y_i)_{i=1}^{\infty}$ 满足

$$\sum_{i=1}^{\infty}|x_i|^p<\infty,\quad \sum_{i=1}^{\infty}|y_i|^p<\infty,$$

则级数 $\sum_{i=1}^{\infty}|x_i+y_i|^p$ 收敛, 而且成立以下的 Minkowski 不等式[5)]:

$$\left(\sum_{i=1}^{\infty}|x_i+y_i|^p\right)^{\frac{1}{p}}\leqslant\left(\sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}}+\left(\sum_{i=1}^{\infty}|y_i|^p\right)^{\frac{1}{p}}.$$

证明 (i) 一个简单的不等式: 如果 $p>1$, $\frac{1}{p}+\frac{1}{q}=1$, 则对任何 $\alpha>0$, $\beta>0$, 成立

$$\alpha\beta\leqslant\frac{\alpha^p}{p}+\frac{\beta^q}{q}.$$

为证这一不等式, 注意到由指数函数的凸性可知, 对任何 $0<\theta<1$, $r\in\mathbb{R}$, $s\in\mathbb{R}$, 有

$$e^{\theta r+(1-\theta)s}\leqslant\theta e^r+(1-\theta)e^s.$$

在上述不等式中取 $\theta=\frac{1}{p}, r=p\mathrm{Log}\,\alpha$, $s=q\mathrm{Log}\,\beta$ 即得要证的不等式.

(ii) Hölder 不等式. 设 $x\neq 0$, $y\neq 0$ (否则, Hölder 不等式显然成立), 又设 $\|x\|_p:=(\sum_{i=1}^{\infty}|x_i|^p)^{\frac{1}{p}}, \|y\|_q:=(\sum_{i=1}^{\infty}|y_i|^q)^{\frac{1}{q}}$ (这里, $\|x\|_p$ 和 $\|y\|_q$ 系便于使用的简略记号). 在 (i) 的不等式中令 $\alpha=\frac{|x_i|}{\|x\|_p}$, $\beta=\frac{|y_i|}{\|y\|_q}$ 可得对每个正整数 i,

$$\frac{|x_iy_i|}{\|x\|_p\|y\|_q}\leqslant\frac{|x_i|^p}{p(\|x\|_p)^p}+\frac{|y_i|^q}{q(\|y\|_q)^q},$$

因而对任何正整数 n,

$$\frac{\sum\limits_{i=1}^{n}|x_iy_i|}{\|x\|_p\|y\|_q}\leqslant\frac{\sum\limits_{i=1}^{n}|x_i|^p}{p(\|x\|_p)^p}+\frac{\sum\limits_{i=1}^{n}|y_i|^q}{q(\|y\|_q)^q}\leqslant\frac{1}{p}+\frac{1}{q}=1.$$

4) O. Hölder [1889]: Über einen Mittelwertsatz. Göttinger Nachrichten, 38–47.

5) H. Minkowski [1896]: Geometrie der Zahlen. Leipzig.

令 $n \to \infty$ 取极限即得级数 $\sum_{i=1}^{\infty} |x_i y_i|$ 收敛, 且 Hölder 不等式成立.

(iii) Minkowski 不等式. 设 $p > 1$ (当 $p = 1$ 时 Minkowski 不等式显然成立), 又取 q 满足 $\frac{1}{p} + \frac{1}{q} = 1$, 于是 $pq - q = p$. 由 Hölder 不等式 (见 (ii)), 对每个正整数 n,

$$
\begin{aligned}
&\sum_{i=1}^{\infty}(|x_i| + |y_i|)^p \\
&= \sum_{i=1}^{n} |x_i|(|x_i| + |y_i|)^{p-1} + \sum_{i=1}^{\infty} |y_i|(|x_i| + |y_i|)^{p-1} \\
&\leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n}(|x_i| + |y_i|)^p\right)^{\frac{1}{q}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n}(|x_i| + |y_i|)^p\right)^{\frac{1}{q}} \\
&= \left(\left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}\right) \left(\sum_{i=1}^{n}(|x_i| + |y_i|)^p\right)^{\frac{1}{q}}.
\end{aligned}
$$

因为 $1 - \frac{1}{q} = \frac{1}{p}$, 由上述不等式即得

$$
\begin{aligned}
\left(\sum_{i=1}^{n} |x_i + y_i|^p\right)^{\frac{1}{p}} &\leqslant \left(\sum_{i=1}^{n}(|x_i| + |y_i|)^p\right)^{\frac{1}{p}} \\
&\leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} |y_i|^p\right)^{\frac{1}{p}}.
\end{aligned}
$$

在上述不等式中, 先令右边的 $n \to \infty$, 再令左边 $n \to \infty$, 即得 $\sum_{i=1}^{n} |x_i + y_i|^p$ 收敛, 而且 Minkowski 不等式成立. □

现在, 我们来定义实或复的赋范向量空间

$$(\ell^p, \|\cdot\|_p), \quad 1 \leqslant p \leqslant \infty,$$

它们就是空间 $(\mathbb{K}^n, \|\cdot\|_p)$ (定理 2.2-2) 到无限序列空间的推广[6)]. 我们还将说明, 除 $p = \infty$ 的情况而外, 这些空间是可分的.

定理 2.4-2 对每个扩充实数 $1 \leqslant p \leqslant \infty$, 用 ℓ^p 表示由 $x_i \in \mathbb{K}$ 的满足下列条件的无穷序列 $x = (x_i)_{i=1}^{\infty}$ 组成的集合, 其中当 $1 \leqslant p < \infty$ 时

$$\sum_{i=1}^{\infty} |x_i|^p < \infty,$$

[6)] 空间 ℓ^p 和 $L^p(\Omega)$ 是由 Frigyes Riesz (1880—1956) 在 1910 年引入的, 他对泛函分析做出了许多里程碑式的以其命名的贡献 (对此将在本章和下一章中做详细的介绍). 他和学生 Béla Szökefalvi Nagy (1913—1998) 合著的 Riesz & Nagy [1955] 是泛函分析方面有极大影响的一部名著.

Frigyes Riesz 有一个兄弟 Marcel Riesz (1886—1969), 也是一位著名的数学家.

当 $p=\infty$ 时

$$\sup_{i\geqslant 1}|x_i|<\infty.$$

(a) 对每个 $1\leqslant p\leqslant\infty$, ℓ^p 为向量空间, 映射 $\|\cdot\|_p$ 定义为当 $1\leqslant p<\infty$ 时

$$x=(x_i)_{i=1}^{\infty}\in\ell^p\to\|x\|_p=\left(\sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}},$$

当 $p=\infty$ 时

$$x=(x_i)_{i=1}^{\infty}\in\ell^{\infty}\to\|x\|_{\infty}=\sup_{i\geqslant 1}|x_i|,$$

则 $\|\cdot\|_p$ 是 ℓ^p 上的范数.

(b) 赋范向量空间 $(\ell^p,\|\cdot\|_p), 1\leqslant p<\infty$, 是可分的.

(c) 赋范向量空间 $(\ell^{\infty},\|\cdot\|_{\infty})$ 是不可分的.

证明 当 $p\geqslant 1$ 时, 由 Minkowski 不等式 (定理 2.4-1), ℓ^p 为向量空间, 当 $p=\infty$ 时, ℓ^p 也显然是向量空间. 类似地, 当 $p\geqslant 1$ 时, $\|\cdot\|_p$ 是 ℓ^p 的范数也可由 Minkowski 不等式导出 (这个不等式即范数的三角不等式, 范数的其他性质可直接验证). 当 $p=\infty$ 时, $\|\cdot\|_{\infty}$ 为范数是显然的. 于是 (a) 得证.

给定 $1\leqslant p<\infty$, 当 $\mathbb{K}=\mathbb{R}$ 时, 令

$$A:=\bigcup_{n=1}^{\infty}\{(y_i)_{i=1}^{\infty}\in\ell^p;\ 当\ i\leqslant n\ 时\ y_i\in\mathbb{Q},\ 当\ i\geqslant n+1\ 时\ y_i=0\},$$

当 $\mathbb{K}=\mathbb{C}$ 时, 令

$$A:=\bigcup_{n=1}^{\infty}\{(y_i)_{i=1}^{\infty}\in\ell^p;\ 当\ i\leqslant n\ 时\ \operatorname{Re}y_i\in\mathbb{Q},\ 且\ \operatorname{Im}y_i\in\mathbb{Q},\ 当\ i\geqslant n+1\ 时\ y_i=0\}.$$

于是, 集合 A 作为可列个可列集的并集, 仍是可列集. 而且 A 在 ℓ^p 中稠密: 这是因为对任何 $x=(x_i)_{i=1}^{\infty}\in\ell^p$ 和任意的 $\varepsilon>0$, 存在 $n_0=n_0(x,\varepsilon)\geqslant 1$, 使得 $\sum_{i=n_0+1}^{\infty}|x_i|^p\leqslant\frac{\varepsilon^p}{2}$. 对于 $1\leqslant i\leqslant n_0$, 当 $\mathbb{K}=\mathbb{R}$ 时, 取 $y_i\in\mathbb{Q}$, 当 $\mathbb{K}=\mathbb{C}$ 时, 取 $y_i\in\mathbb{C}$, 使 $\operatorname{Re}y_i\in\mathbb{Q}$, 且 $\operatorname{Im}y_i\in\mathbb{Q}$, 要求 $\sum_{i=1}^{n_0}|x_i-y_i|^p\leqslant\frac{\varepsilon^p}{2}$. 这样, 向量 $y:=(y_1,\cdots,y_{n_0},0,\cdots)$ 属于 A, 且满足 $\|y-x\|_p\leqslant\varepsilon$. 于是 (b) 得证.

考察集合

$$B:=\{(x_i)_{i=1}^{\infty}\in\ell^{\infty};\ 当\ i\geqslant 1\ 时\ x_i=0\ 或\ x_i=1\}.$$

因为 $(x_i)_{i=1}^{\infty}\in B\to\sum_{i=1}^{\infty}\frac{x_i}{2^i}$ 是 B 到 $[0,1[$ 上的一个满射 (参见 1.5 节, 利用区间 $[0,1[$ 中每个实数均有二进展开的性质), 因而 B 是 ℓ^{∞} 的一个不可列子集.

设 C 是 ℓ^{∞} 中任一稠密子集. 任意给定 $x\in B$, 必定存在 $y(x)\in C$, 使得 $\|x-y(x)\|_{\infty}<\frac{1}{2}$, 映射 $x\in B\to y(x)\in C$ 必为单射 (因为 $x,\tilde{x}\in B$ 且当 $x\neq\tilde{x}$ 时, $\|x-\tilde{x}\|_{\infty}=1$, 从而 $y(x)\neq y(\tilde{x})$), 所以 C 必为不可列集 (1.5 节). 于是 (c) 得证. □

定理 2.4-1 和 2.4-2 的若干有趣补充可见下面的练习.

习题

2.4-1 (1) 给定任意向量 $x \in \mathbb{K}^n$, 证明: $\|x\|_\infty = \lim_{p\to\infty} \|x\|_p$.

(2) 给定任意的 $x = (x_i)_{i=1}^\infty \in \ell^\infty$, 证明: $\|x\|_\infty = \lim_{n\to\infty}\{\lim_{p\to\infty}(\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}\}$.

2.4-2 (1) 证明 Hölder 不等式 (定理 2.4-1) 中等式成立当且仅当存在常数 $\alpha \geqslant 0$, $\beta \geqslant 0$, 使 $\alpha + \beta > 0$, 满足对所有的 $i \geqslant 1$ 有 $\alpha|x_i|^p = \beta|y_i|^q$.

(2) 证明 Minkowski 不等式 (定理 2.4-1) 中等式成立当且仅当存在常数 $\alpha \geqslant 0$, $\beta \geqslant 0$, 使 $\alpha + \beta > 0$, 满足对所有的 $i \geqslant 1$, 有 $\alpha x_i = \beta y_i$.

2.4-3 给定实数 $0 < p < 1$, 设 X 为所有满足 $\sum_{i=1}^\infty |x_i|^p < \infty$ 的数列 $(x_i)_{i=1}^\infty$ 的集合.

(1) 证明 X 为向量空间.

(2) 证明映射 $(x_i)_{i=1}^\infty \in X \to (\sum_{i=1}^\infty |x_i|^p)^{\frac{1}{p}}$ 不是 X 上的范数.

(3) 定义映射 $d : X \times X \to \mathbb{R}$ 为 $d(x, y) = \sum_{i=1}^\infty |x_i - y_i|^p$, 其中 $x = (x_i)_{i=1}^\infty \in X, y = (y_i)_{i=1}^\infty \in X$, 证明 d 是 X 上的距离.

2.4-4 设 p, q 为两个实数, 满足 $0 < p < q$, 数 $x_i \in \mathbb{K}$ 的无穷序列 $(x_i)_{i=1}^\infty$ 满足 $\sum_{i=1}^\infty |x_i|^p < \infty$. 证明级数 $\sum_{i=1}^\infty |x_i|^q$ 也收敛, 且 ℓ^p 中的 Jensen 不等式[7]成立:

$$\left(\sum_{i=1}^\infty |x_i|^q\right)^{\frac{1}{q}} \leqslant \left(\sum_{i=1}^\infty |x_i|^p\right)^{\frac{1}{p}}.$$

注意, 由 Jensen 不等式, 对每个 $p \geqslant 1$, 空间 ℓ^p 包含于所有的空间 ℓ^q 中, 其中 $p < q \leqslant \infty$, 而且对任何 $x \in \ell^p$, 有 $\|x\|_q \leqslant \|x\|_p$.

2.5 Lebesgue 空间 $L^p(\Omega)$, $1 \leqslant p \leqslant \infty$

在本节中, 用 Ω 表示 $\mathbb{R}^n$ 中任意一个开 (从而可测) 子集. 相应的空间 $L^1(\Omega)$ (对 $\mathbb{R}^n$ 中任意的可测子集 A, 向量空间 $L^1(A)$ 的定义及主要性质参见 1.15 节) 由所有的实 Lebesgue 可积函数 (的等价类) 组成, 即其元素为可测函数 $f : \Omega \to [-\infty, \infty]$, 满足

$$\int_\Omega |f(x)| dx < \infty.$$

现在, 扩充一下这个定义. 给定任何 $1 < p < \infty$, 用 $L^p(\Omega)$ 表示所有的可测函数 (的等价类) $f : \Omega \to [-\infty, \infty]$, 使得 $|f|^p \in L^1(\Omega)$, 或等价地满足

$$\int_\Omega |f(x)|^p dx < \infty$$

组成的集合.

---

[7] J. L. W. V. Jensen [1906]: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica 30, 175–193.

本节的第一个目的是证明这样定义的集合 $L^p(\Omega), 1 \leqslant p < \infty$ 和后面 (定理 2.5-2) 将要定义的集合 $L^\infty(\Omega)$ 均为 (实) 赋范向量空间. 为此, 我们将沿着讨论赋范向量空间 $(\ell^p, \|\cdot\|_p), 1 \leqslant p \leqslant \infty$ (2.4 节) 类似的途径, 读者可将定理 2.5-1 和定理 2.5-2 与定理 2.4-1 和定理 2.4-2 (a) 的叙述与证明分别作比较.

定理 2.5-1 (函数的 Hölolor 不等式和 Minkowski 不等式) 令 Ω 为 $\mathbb{R}^n$ 的开子集.

(a) 给定实数 $p > 1$, 实数 q 满足 $\frac{1}{p} + \frac{1}{q} = 1$ (从而 $q > 1$), 设 $f : \Omega \to [-\infty, \infty]$ 和 $g : \Omega \to [-\infty, \infty]$ 为两个可测函数, 满足

$$\int_\Omega |f(x)|^p dx < \infty, \quad \int_\Omega |g(x)|^q dx < \infty,$$

则 $fg \in L^1(\Omega)$, 而且成立 Hölder 不等式:

$$\int_\Omega |f(x)g(x)| dx \leqslant \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}} \left(\int_\Omega |g(x)|^q dx\right)^{\frac{1}{q}}.$$

(b) 给定实数 $p \geqslant 1$. 设 $f : \Omega \to [-\infty, \infty]$ 和 $g : \Omega \to [-\infty, \infty]$ 为两个可测函数, 满足

$$\int_\Omega |f(x)|^p dx < \infty, \quad \int_\Omega |g(x)|^p dx < \infty,$$

则 $f + g \in L^p(\Omega)$, 而且成立 Minkowski 不等式:

$$\left(\int_\Omega |f(x) + g(x)|^p dx\right)^{\frac{1}{p}} \leqslant \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}} + \left(\int_\Omega |g(x)|^p dx\right)^{\frac{1}{p}}.$$

证明 (i) Hölder 不等式. 设 $f \neq 0, g \neq 0$ (否则, Hölder 不等式显然成立), 设 $\|f\|_p := \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}}, \|g(x)\|_q := \left(\int_\Omega |g(x)|^q dx\right)^{\frac{1}{q}}$(这里, $\|f\|_p$ 和 $\|g\|_q$ 系便于使用的简略记号). 令 $\alpha := \frac{|f(x)|}{\|f\|_p}, \beta := \frac{|g(x)|}{\|g\|_q}$, 由不等式 $\alpha\beta \leqslant \frac{\alpha^p}{p} + \frac{\beta^q}{q}$ (见定理 2.4-1 证明中的 (i)) 可得, 对所有的 $x \in \Omega$,

$$\frac{|f(x)g(x)|}{\|f\|_p \|g\|_q} \leqslant \frac{1}{p} \frac{|f(x)|^p}{(\|f\|_p)^p} + \frac{1}{q} \frac{|g(x)|^q}{(\|g\|_q)^q}.$$

所以

$$\begin{aligned} \frac{1}{\|f\|_p \|g\|_q} \int_\Omega |f(x)g(x)| dx &\leqslant \frac{1}{p(\|f\|_p)^p} \int_\Omega |f(x)|^p dx + \frac{1}{q(\|g\|_q)^q} \int_\Omega |g(x)|^q dx \\ &= \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

因此, $fg \in L^1(\Omega)$, 而且 Hölder 不等式成立. (a) 得证.

(ii) Minkowski 不等式. 设 $p > 1$ (当 $p = 1$ 时 Minkowski 不等式显然成立), 又设 q 满足 $\frac{1}{p} + \frac{1}{q} = 1$, 这样 $pq - q = p$. 由 Hölder 不等式 (第 (i) 部分) 可得

$$
\begin{aligned}
&\int_\Omega (|f(x)| + |g(x)|)^p dx \\
&= \int_\Omega |f(x)|(|f(x)| + |g(x)|)^{p-1} dx + \int_\Omega |g(x)|(|f(x)| + |g(x)|)^{p-1} dx \\
&\leqslant \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}} \left(\int_\Omega (|f(x)| + |g(x)|)^p dx\right)^{\frac{1}{q}} \\
&\quad + \left(\int_\Omega |g(x)|^p dx\right)^{\frac{1}{p}} \left(\int_\Omega (|f(x)| + |g(x)|)^p dx\right)^{\frac{1}{q}} \\
&= \left\{\left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}} + \left(\int_\Omega |g(x)|^p dx\right)^{\frac{1}{p}}\right\} \left(\int_\Omega (|f(x)| + |g(x)|)^p dx\right)^{\frac{1}{q}}.
\end{aligned}
$$

因为 $1 - \frac{1}{q} = \frac{1}{p}$, 由上述不等式即得

$$
\begin{aligned}
\left(\int_\Omega |f(x) + g(x)|^p dx\right)^{\frac{1}{p}} &\leqslant \left(\int_\Omega (|f(x)| + |g(x)|)^p dx\right)^{\frac{1}{p}} \\
&\leqslant \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}} + \left(\int_\Omega |g(x)|^p dx\right)^{\frac{1}{p}}.
\end{aligned}
$$

从而 $f + g \in L^p(\Omega)$, 而且 Minkowski 不等式成立. (b) 得证. □

现在我们来定义赋范向量空间

$$
\left(L^p(\Omega), \|\cdot\|_{L^p(\Omega)}\right), \ 1 \leqslant p \leqslant \infty,
$$

它们被称为 Lebesgue 空间[8].

定理 2.5-2 对每个扩充实数 $1 \leqslant p \leqslant \infty$, 用 $L^p(\Omega)$ 表示满足下列条件的可测函数 $f : \Omega \to [-\infty, \infty]$ 的集合, 当 $1 \leqslant p < \infty$ 时

$$
\int_\Omega |f(x)|^p dx < \infty,
$$

当 $p = \infty$ 时

$$
\inf\{C \geqslant 0;\ \text{在 } \Omega \text{ 上几乎处处有 } |f| \leqslant C\} < \infty.
$$

定义映射 $\|\cdot\|_{L^p(\Omega)}$ 如下:

$$
\begin{aligned}
&f \in L^p(\Omega) \to \|f\|_{L^p(\Omega)} := \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}}, \quad 1 \leqslant p < \infty, \\
&f \in L^\infty(\Omega) \to \|f\|_{L^\infty(\Omega)} := \inf\{C \geqslant 0;\ \text{在 } \Omega \text{ 上几乎处处 } |f| \leqslant C\}.
\end{aligned}
$$

[8] 这样命名是为了纪念 Lebesgue 测度和 Lebesgue 积分理论的创始人 Henri Lebesgue (1875—1941), 和他的 Note aux Comptes Rendus:

H. Lebesgue [1901]: Sur une généralisation de l'intégrale définie. Comptes Rendus des Séances de l'Académie des Sciences 132, 1025–1027.

则对每个 $1 \leqslant p \leqslant \infty$, $L^p(\Omega)$ 为向量空间, $\|\cdot\|_{L^p(\Omega)}$ 是 $L^p(\Omega)$ 上的范数.

证明 对 $1 \leqslant p < \infty$, 由 Minkowski 不等式 (定理 2.5-1) 可知 $L^p(\Omega)$ 是向量空间, $\|\cdot\|_{L^p(\Omega)}$ 是 $L^p(\Omega)$ 上的一个范数 (Minkowski 不等式即 $\|\cdot\|_{L^p(\Omega)}$ 的三角不等式, 范数的其他性质则容易直接验证). 又显然 $L^\infty(\Omega)$ 为向量空间, $\|\cdot\|_{L^\infty(\Omega)}$ 是 $L^\infty(\Omega)$ 上的一个范数. □

注 对于 $1 \leqslant p < \infty$, 类似地可定义复空间

$$L^p(\Omega;\mathbb{C}) := \{f : \Omega \to \mathbb{C};\ \mathrm{Re}\, f \text{ 和 } \mathrm{Im}\, f \text{ 均可测, 且 } |f|^p \in L^1(\Omega)\},$$

它们具有空间 $L^p(\Omega)$ 的类似性质[9].

给定可测函数 $f : \Omega \to [-\infty, \infty]$, 称扩充实数

$$\inf\{C \geqslant 0;\ \text{在 } \Omega \text{ 上几乎处处 } |f| \leqslant C\} \in [0, \infty]$$

为 f 的本性最大模. 空间 $L^\infty(\Omega)$ 即是由所有具有有限本性最大模的可测函数 (的等价类) 组成.

关于空间的可分性, 在 ℓ^p 空间情况下讨论相对容易 (定理 2.4-2(b)), 而对 Lebesgue 空间 $L^p(\Omega)$ 却不再如此. 作为 $1 \leqslant p < \infty$ 情况下的初步结果, 我们将首先在定理 2.5-3 中证明: Lebesgue 空间 $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)}), 1 \leqslant p < \infty$ 中的任何函数, 可以用在 Ω 中具有紧支集的连续函数任意地逼近. 这个有趣的结果将在定理 2.6-2 中进一步深化, 那里我们将证明 $L^p(\Omega), 1 \leqslant p < \infty$ 中的任何函数, 可以用具有在 Ω 中紧支集的无限阶可微函数任意地逼近.

注意下面的结果对于 $p = \infty$ 并不成立, 参见习题 2.5-5.

定理 2.5-3 设 Ω 是 $\mathbb{R}^n$ 中的开子集, 空间

$$\mathcal{C}_c(\Omega) = \{g \in \mathcal{C}(\Omega); \mathrm{supp}\, g \text{ 是 } \Omega \text{ 的紧子集}\}.$$

则对每个 $1 \leqslant p < \infty$, 子空间 $\mathcal{C}_c(\Omega)$ 在 $L^p(\Omega)$ 中稠密.

证明 设 $f \in L^p(\Omega)$, 任意给定 $\varepsilon > 0$. 我们的目标是找到函数 $g \in \mathcal{C}_c(\Omega)$, 使得 $\|f - g\|_{L^p(\Omega)} \leqslant \varepsilon$.

(i) 存在可测的简单函数 $s = s(f, \varepsilon)$, 使得

$$\mu(\{x \in \Omega; s(x) \neq 0\}) < \infty \text{ 且 } \|f - s\|_{L^p(\Omega)} < \varepsilon/2,$$

其中 μ 表示 $\mathbb{R}^n$ 上的 Lebesgue 测度, 可测简单函数的定义已在 1.14 节给出.

为证这个结论, 先设 $f \geqslant 0$. 于是, 由定理 1.14-5, 存在一列具有下述性质的可测简单函数 $(s_k)_{k=1}^\infty$:

[9] 对空间 $L^p(\Omega;\mathbb{C})$, 在 Hewitt 和 Stromberg [1965, 13 节] 中有详尽的分析.

对所有的 $k \geqslant 1$, $0 \leqslant s_k \leqslant f$; 对每个 $x \in \Omega$, 当 $k \to \infty$ 时 $s_k(x) \to f(x)$.

因此, 对所有的 $k \geqslant 1$, $s_k \in L^p(\Omega)$, 故 $\mu(\{x \in \Omega; s_k(x) \neq 0\}) < \infty$. 且对每个 $x \in \Omega$,

$$|(f - s_k)(x)|^p \leqslant |f(x)|^p, \quad k \geqslant 1,$$
$$|(f - s_k)(x)|^p \to 0, \quad k \to \infty.$$

将 Lebesgue 控制收敛定理 (定理 1.15-3) 应用于函数 $|f - s_k|^p \in L^1(\Omega), k \geqslant 1$, 由于它们被同一个函数 $|f|^p \in L^1(\Omega)$ 所控制, 从而当 $k \to \infty$ 时

$$\int_\Omega |f(x) - s_k(x)|^p dx \to 0.$$

回到一般的情况. 令

$$\Omega^+ := \{x \in \Omega; f(x) > 0\}, \quad \Omega^- := \{x \in \Omega; f(x) < 0\}.$$

由上述论证可知, 存在可测的简单函数 $s_k^+ : \Omega^+ \to [0, \infty[$, $s_k^- : \Omega^- \to [0, \infty[$, $k \geqslant 1$, 使得当 $k \to \infty$ 时

$$\int_{\Omega^+} |f(x) - s_k^+(x)|^p dx \to 0,$$
$$\int_{\Omega^-} |-f(x) - s_k^-(x)|^p dx \to 0.$$

对每个 $k \geqslant 1$, 定义可测简单函数 $s_k : \Omega \to \mathbb{R}$ 如下: 在 Ω^+ 上, $s_k := s_k^+$, 在 Ω^- 上, $s_k := s_k^-$, 在 $\Omega - (\Omega^+ \cup \Omega^-)$ 上, $s_k := 0$. 于是, 当 $k \to \infty$ 时

$$\begin{aligned}&\int_\Omega |f(x) - s_k(x)|^p dx \\ &= \int_{\Omega^+} |f(x) - s_k^+(x)|^p dx + \int_{\Omega^-} |-f(x) - s_k^-(x)|^p dx \to 0.\end{aligned}$$

因而 (i) 得证.

(ii) 设 $s = s(f, \varepsilon)$ 为由 (i) 构造的可测简单函数, 则存在函数 $g = g(s, \varepsilon) = g(f, \varepsilon) \in \mathcal{C}_c(\Omega)$, 使得

$$\|s - g\|_{L^p(\Omega)} \leqslant \frac{\varepsilon}{2}.$$

因为 $\mu(\{x \in \Omega; s(x) \neq 0\}) < \infty$, 由 Lusin 性质 (定理 1.14-4 (c)), 存在函数 $g \in \mathcal{C}_c(\Omega)$, 使得

$$\sup_{x \in \Omega} |g(x)| \leqslant \|s\|_{L^\infty(\Omega)}, \mu(\{x \in \Omega; g(x) \neq s(x)\}) \leqslant \left(\frac{\varepsilon}{4\|s\|_{L^\infty(\Omega)}}\right)^p.$$

因为对一切 $x \in \Omega, |s(x) - g(x)| \leqslant 2\|s\|_{L^\infty}$, 所以

$$\|s - g\|_{L^p(\Omega)} = \left(\int_{\{x\in\Omega; g(x)\neq s(x)\}} |s(x) - g(x)|^p dx \right)^{\frac{1}{p}} \leqslant \frac{\varepsilon}{2}.$$

这样便证得 (ii) □

注意, 由定理 2.5-3 及其证明中的部分 (i), 可以得到定义 Lebesgue 空间 $L^p(\Omega)$, $1 \leqslant p < \infty$ 的另两条途径, 即或者可把它作为空间 $\mathcal{C}_c(\Omega)$ 按范数 $\|\cdot\|_{L^p(\Omega)}$ 的完备化 (在这个定义中可采用 Riemann 积分), 或者也可把它作为由所有满足 $\int_\Omega |s|^p dx < \infty$ 的可测简单函数 $s : \Omega \to \mathbb{R}$ 组成的空间按范数 $\|\cdot\|_{L^p(\Omega)}$ 的完备化.

还可注意, 由定理 2.5-3 可知, 如果 Ω 有界, 则空间 $\mathcal{C}(\overline{\Omega})$ 在 $L^p(\Omega)$ 中稠密, 此处 $1 \leqslant p < \infty$.

现在, 我们就来讨论空间 $L^p(\Omega)$ 的可分性.

定理 2.5-4 设 Ω 为 $\mathbb{R}^n$ 中开子集.

(a) 赋范向量空间 $L^p(\Omega)$, $1 \leqslant p < \infty$, 是可分空间.

(b) 赋范向量空间 $L^\infty(\Omega)$ 是不可分的.

证明 下面用 χ_A 表示集合 A 的特征函数.

(i) 首先, 设 $f \in L^p(\Omega)$, $1 \leqslant p < \infty$, 任意给定 $\varepsilon > 0$. 由定理 2.5-3, 存在函数 $g = g(f, \varepsilon) \in \mathcal{C}_c(\Omega)$, 使得

$$\|f - g\|_{L^p(\Omega)} \leqslant \frac{\varepsilon}{2}.$$

集合 $K := \operatorname{supp} g$ 是 Ω 的一个紧子集, 存在有界开集 U, 使得 $K \subset U \subset \Omega$ (为此, 只要考察以 K 中的点为中心的开球构成的 K 的一个覆盖, 取 U 为 K 的一个有限子覆盖即可).

连续函数 g 在紧集 $\overline{U}$ 上一致连续 (定理 1.13-2), 因而存在 $\delta_0 > 0$, 使得任何 $x, y \in \overline{U}$, 且当 $\|x - y\|_\infty < \delta_0$ 时

$$|g(x) - g(y)| \leqslant \tilde{\varepsilon} := \frac{\varepsilon}{2(\mu(U))^{\frac{1}{p}}},$$

其中 μ 为 $\mathbb{R}^n$ 中的 Lebesgue 测度. 又由函数 $x \in K \to \inf_{y\in\mathbb{R}^n - U} \|x - y\|_\infty$ 的连续性 (定理 1.11-3) 和 K 的紧性, 存在 $\delta_1 > 0$, 使得对任何 $x \in K$,

$$\{y \in \mathbb{R}^n; \|y - x\|_\infty < \delta_1\} \subset U.$$

取 $\delta \in \mathbb{Q}$, 使得 $0 < \delta < \min\{\delta_0, \delta_1\}$.

记所有形如

$$\{y \in \mathbb{R}^n; \|y - x\|_\infty < \frac{\delta}{2}, \text{ 其中 } x_j = p_j\delta, p_j \in \mathbb{Z}, 1 \leqslant j \leqslant n\}$$

的开球组成的可数无限族为 $(B_i)_{i\in I}$. 用 $(B_i)_{i\in I(K)}$ 表示满足 $B_i \cap K \neq \varnothing$ 的球 $B_i, i \in I$, 组成的子族. 于是, 对每个 $i \in I(K)$, 存在 $\alpha_i \in \mathbb{Q}$, 使得对所有的 $y \in B_i$ 有

$$|g(y) - \alpha_i| \leqslant \tilde{\varepsilon}.$$

(如果函数 $g|_{\overline{B}_i}$ 并非常数, 可在其最小值和最大值之间任取 $\alpha_i \in \mathbb{Q}$; 如果 $g|_{\overline{B}_i}$ 为常数 β_i, 则任取 $\alpha_i \in \mathbb{Q}$ 使得 $|\alpha_i - \beta_i| \leqslant \tilde{\varepsilon}$.) 这样, 考察函数

$$h := \sum_{i\in I(K)} \alpha_i \chi_{B_i},$$

由构造, 对几乎所有的 $x \in U$, 满足 $|h(x) - g(x)| \leqslant \tilde{\varepsilon}$, 所以

$$\begin{aligned}\|h - g\|_{L^p(\Omega)} &= \left(\int_U |h(x) - g(x)|^p dx\right)^{\frac{1}{p}}\\ &\leqslant (\mu(U))^{\frac{1}{p}}\tilde{\varepsilon} = \frac{\varepsilon}{2}.\end{aligned}$$

由此可得 $\|f - h\|_{L^p(\Omega)} \leqslant \dfrac{\varepsilon}{2}$. 再注意到这样的 h 构成一个可数无限族 (因为对所有的 $i \in I(K), \alpha_i \in \mathbb{Q}$, 集合 $I(K)$ 是可数无限的), 所以空间 $L^p(\Omega), 1 \leqslant p < \infty$ 是可分的.

(ii) 其次, 设 $p = \infty$. 给定任意的 $x \in \Omega$, 设 $B(x)$ 为以 x 为中心且包含于 Ω 中的任意开球, 又设

$$O(x) := \left\{f \in L^\infty(\Omega); \|f - \chi_{B(x)}\|_{L^\infty(\Omega)} < \frac{1}{2}\right\}.$$

于是 $(O(x))_{x\in\Omega}$ 是 $L^\infty(\Omega)$ 的非空开子集组成的不可数无限族, 且满足当 $x \neq y$ 时

$$O(x) \cap O(y) = \varnothing.$$

(如果 $x \neq y$, 存在开球 B 使得 $B \subset B(x)$ 且 $B \cap B(y) = \varnothing$, 但是对于 $b \in B$, 不等式 $|f(b) - 1| < \frac{1}{2}$ 和 $|f(b)| < \frac{1}{2}$ 不可能同时成立.)

假设 $L^\infty(\Omega)$ 的可数子集 $\cup_{n\in\mathbb{N}}\{f_n\}$ 在 $L^\infty(\Omega)$ 中稠密. 于是, 对每个 $x \in \Omega$, 存在整数 $n(x) \geqslant 0$, 使 $f_{n(x)} \in O(x)$. 这样定义的映射 $x \in \Omega \to n(x) \in \mathbb{N}$ 必定是单射 (当 $x \neq y$ 时 $f_{n(x)} \in O(x), f_{n(y)} \in O(y)$, 且 $O(x) \cap O(y) = \varnothing$), 这就导致矛盾 (1.5 节). □

注 定理 2.5-4(a) 的另一个简单的证明可利用多变量函数的 Weierstrass 逼近定理 (定理 2.15-2) 导出, 参见习题 2.15-2.

习题

在下面题目中, Ω 表示 $\mathbb{R}^n$ 中的开集.

2.5-1 (1) 证明 Hölder 不等式 (定理 2.5-1(a)) 中等式成立当且仅当存在常数 $\alpha \geqslant 0, \beta \geqslant 0, \alpha+\beta>0$, 使得对几乎所有的 $x \in \Omega, \alpha|f(x)|^p=\beta|g(x)|^q$.

(2) 证明 Minkowski 不等式 (定理 2.5-1(b)) 中等式成立, 当 $p=1$ 时当且仅当存在可测函数 $h: \Omega \rightarrow [0, \infty[$, 使得 $f=gh$ 在集合 $\{x \in \Omega; f(x) \neq 0$ 且 $g(x) \neq 0\}$ 上几乎处处成立; 当 $1<p<\infty$ 时当且仅当存在常数 $\alpha \geqslant 0, \beta \geqslant 0, \alpha+\beta>0$, 使得对几乎所有的 $x \in \Omega, \alpha f(x)=\beta g(x)$.

2.5-2 给定 $1 \leqslant p<\infty$, 设函数 $f_k \in L^p(\Omega), k \geqslant 1, f \in L^p(\Omega)$ 满足当 $k \rightarrow \infty$ 时, 在 Ω 上 $(f_k)_{k=1}^{\infty}$ 几乎处处收敛于 f, 且

$$\|f_k\|_{L^p(\Omega)} \rightarrow \|f\|_{L^p(\Omega)}.$$

证明: 当 $k \rightarrow \infty$ 时 $\|f_k-f\|_{L^p(\Omega)} \rightarrow 0$.

2.5-3 (1) 给定 $0<p<1$. 设 q 定义为 $\frac{1}{p}+\frac{1}{q}=1$ (因而 $q<0$). 设 $f: \Omega \rightarrow [-\infty, \infty], g: \Omega \rightarrow [-\infty, \infty]$ 为两个可测函数, 满足

$$\int_{\Omega}|f(x)|^p dx<\infty, \quad 0<\int_{\Omega}|g(x)|^q dx<\infty.$$

证明下述反向 Hölder 不等式:

$$\int_{\Omega}|f(x) g(x)| dx \geqslant\left(\int_{\Omega}|f(x)|^p dx\right)^{\frac{1}{p}}\left(\int_{\Omega}|g(x)|^q dx\right)^{\frac{1}{q}},$$

这个不等式的左边可能为 ∞.

(2) 给定 $0<p<1$. 设 $f: \Omega \rightarrow [-\infty, \infty], g: \Omega \rightarrow [-\infty, \infty]$ 为两个可测函数, 满足

$$\int_{\Omega}|f(x)|^p dx<\infty, \quad \int_{\Omega}|g(x)|^p dx<\infty.$$

证明下述反向 Minkowski 不等式:

$$\left(\int_{\Omega}(|f(x)|+|g(x)|)^p dx\right)^{\frac{1}{p}} \geqslant\left(\int_{\Omega}|f(x)|^p dx\right)^{\frac{1}{p}}+\left(\int_{\Omega}|g(x)|^p dx\right)^{\frac{1}{p}}.$$

2.5-4 给定 $0<p<1$. 用 $L^p(\Omega)$ 表示所有满足 $\int_{\Omega}|f(x)|^p dx<\infty$ 的可测函数 $f: \Omega \rightarrow [-\infty, \infty]$ 的集合.

(1) 证明 $L^p(\Omega)$ 为向量空间.

提示: 证明对任何 $f, g \in L^p(\Omega)$,

$$\int_{\Omega}|f(x)+g(x)|^p dx \leqslant 2^{1-p}\left\{\left(\int_{\Omega}|f(x)|^p dx\right)^{\frac{1}{p}}+\left(\int_{\Omega}|g(x)|^p dx\right)^{\frac{1}{p}}\right\}^p.$$

(2) 映射 $d_p: L^p(\Omega) \times L^p(\Omega) \rightarrow [0, \infty[$ 定义为

$$d_p(f, g)=\int_{\Omega}|f(x)-g(x)|^p dx,$$

其中 $f, g \in L^p(\Omega)$. 证明 d_p 是 $L^p(\Omega)$ 上的一个距离.

2.5-5 证明子空间 $\mathcal{C}_c(\Omega)$ 在 $L^p(\Omega)$ 中不稠密.

2.5-6 设 $1 < p < \infty$.

(1) 证明对于在 $(0, \infty)$ 几乎处处非负的函数 $f \in L^p(0, \infty)$, 成立[10)]

$$\int_0^\infty \left(\frac{1}{x}\int_0^x f(t)dt\right)^p dx \leqslant \left(\frac{p}{p-1}\right)^p \int_0^\infty f(x)^p dx.$$

(2) 证明在这个不等式中常数 $\left(\frac{p}{p-1}\right)^p$ 是最佳的.

(3) 证明不存在非零函数 f 使此不等式中等式成立.

2.6 空间 $L^p(\Omega)$ $(1 \leqslant p < \infty)$ 的正则化与逼近

设 Ω 为 $\mathbb{R}^n$ 中的开子集. 函数 $f : \Omega \to [-\infty, \infty]$ 称为在 Ω 上局部可积的, 是指 f 是可测的, 而且 f 在 Ω 的任何紧子集 K 上的限制 $f|_K$ 属于空间 $\mathcal{L}^1(K)$. 因为 Ω 的任何紧子集均有闭包在 Ω 中紧的开子集的有限覆盖 (如由球组成), 所以可测函数 $f : \Omega \to [-\infty, \infty]$ 为局部可积的, 当且仅当对任意给定的 Ω 的开子集 U, 当 $\overline{U}$ 是 Ω 的紧子集时, 其限制 $f|_U$ 在空间 $L^1(U)$ 中.

显然, 局部可积函数组成向量空间, 记作 $\mathcal{L}^1_{\text{loc}}(\Omega)$. 商空间

$$L^1_{\text{loc}}(\Omega) := \mathcal{L}^1_{\text{loc}}(\Omega)/\mathcal{R},$$

其中等价关系 $\mathcal{R}$ 即在 Ω 上几乎处处相等. 显然, $L^1_{\text{loc}}(\Omega)$ 也是一个向量空间. 习惯上, $\mathcal{L}^1_{\text{loc}}(\Omega)$ 中的函数视为等同于它们在 $L^1_{\text{loc}}(\Omega)$ 中的等价类.

注 空间 $L^1_{\text{loc}}(\Omega)$ 上可以赋以距离拓扑, 见习题 2.6-1.

更一般地, 可以定义空间

$$L^p_{\text{loc}}(\Omega), \quad 1 \leqslant p \leqslant \infty.$$

它表示由满足 $f|_U \in L^p(U)$ 的可测函数 $f : \Omega \to [-\infty, \infty]$ 组成的集合, 其中 U 取遍闭包在 Ω 中紧的开子集.

对 $1 \leqslant p \leqslant \infty$, 任何函数 $f \in L^p(\Omega)$ 均在 Ω 上局部可积. 这是因为对 Ω 的任何紧子集 K, 当 $p = 1$ 时

$$\int_K |f(x)|dx \leqslant \|f\|_{L^1(\Omega)} < \infty,$$

[10)] 这个著名的 Hardy 不等式见:

G. H. Hardy [1925]: Notes on some points in the integral calculus. LX. An inequality between integrals, Messengers of Mathematics 54, 150–156.

自此之后, 这个不等式有一系列推广, 概览可见:

A. Kufner; L. Maligranda; L. E. Persson [2007]: The Hardy Inequality: About Its History and some Related Results. Pilsen.

当 $1<p<\infty$ 时, 取 $q=\frac{p}{p-1}$ (由 Hölder 不等式, 见定理 2.5-1(a)),

$$\begin{aligned}\int_K |f(x)|dx &\leqslant \left(\int_K dx\right)^{\frac{1}{q}}\left(\int_K |f(x)|^p dx\right)^{\frac{1}{p}}\\ &\leqslant \left(\int_K dx\right)^{\frac{1}{q}}\|f\|_{L^p(\Omega)}<\infty.\end{aligned}$$

当 $p=\infty$ 时, 取 $q=1$,

$$\int_K |f(x)|dx \leqslant \left(\int_K dx\right)\|f\|_{L^\infty(\Omega)}<\infty.$$

显然, 空间 $\mathcal{C}(\Omega)$ 中任何函数在 Ω 中局部可积, 这是因为对 Ω 的任何紧子集 K,

$$\int_K |f(x)|dx \leqslant \left(\int_K dx\right)\sup_{x\in K}|f(x)|<\infty.$$

$\mathbb{R}^n$ 中一族光滑化子是指由形为

$$\omega_\varepsilon(x)=\frac{1}{\varepsilon^n}\omega\left(\frac{x}{\varepsilon}\right),\quad x\in\mathbb{R}^n,$$

的函数 $\omega_\varepsilon:\mathbb{R}^n\to\mathbb{R}$ 组成的函数族 $(\omega_\varepsilon)_{\varepsilon>0}$, 其中 $\omega:\mathbb{R}^n\to\mathbb{R}$ 满足

$$\begin{gathered}\omega\in\mathcal{C}^\infty(\mathbb{R}^n),\ \text{对所有的}\ x\in\mathbb{R}^n,\ \omega(x)\geqslant 0,\\ \operatorname{supp}\omega\subset\overline{B(0,1)},\quad \int_{\mathbb{R}^n}\omega(x)dx=1.\end{gathered}$$

因此, 对每个 $\varepsilon>0$,

$$\begin{gathered}\omega_\varepsilon\in\mathcal{C}^\infty(\mathbb{R}^n),\ \text{对所有的}\ x\in\mathbb{R}^n,\ \omega_\varepsilon(x)\geqslant 0,\\ \operatorname{supp}\omega_\varepsilon\subset\overline{B(0,\varepsilon)},\quad \int_{\mathbb{R}^n}\omega_\varepsilon(x)dx=1.\end{gathered}$$

这类函数 ω 的例子如

$$\omega(x):=\begin{cases}ce^{\frac{1}{|x|^2-1}}, & |x|<1,\\ 0, & |x|\geqslant 1,\end{cases}$$

其中常数 c 使得 $\int_{B(0,1)}\omega(y)dy=1$ (习题 2.6-2).

设 Ω 为 $\mathbb{R}^n$ 的开子集, 函数 $f\in L^1_{\mathrm{loc}}(\Omega)$, $(\omega_\varepsilon)_{\varepsilon>0}$ 为一族光滑化子, 对于任何 $\varepsilon>0$, 定义集合 Ω_ε 和函数 $f_\varepsilon:\Omega_\varepsilon\to\mathbb{R}$ 如下:

$$\begin{gathered}\Omega_\varepsilon:=\{x\in\Omega;\operatorname{dist}(x,\mathbb{R}^n-\Omega)>\varepsilon\},\\ f_\varepsilon(x):=\int_\Omega \omega_\varepsilon(x-y)f(y)dy,\quad x\in\Omega_\varepsilon.\end{gathered}$$

称族 $(f_\varepsilon)_{\varepsilon>0}$ 为 f 的正则化族.

对每个 $\varepsilon > 0$, 集合 Ω_ε 显然是开集 (函数 $x \in \Omega \to \operatorname{dist}(x, \mathbb{R}^n - \Omega)$ 是连续的, 参见定理 1.11-3). 球 $\overline{B(x;\varepsilon)}$ 包含于 Ω (从而上述函数 f_ε 的定义有意义), 对每个 $x \in \Omega_\varepsilon$, $f_\varepsilon(x)$ 可等价地表示为

$$\begin{aligned} f_\varepsilon(x) &= \int_{B(x;\varepsilon)} \omega_\varepsilon(x-y) f(y) dy = \int_{B(0;\varepsilon)} \omega_\varepsilon(x) f(x-z) dz \\ &= \frac{1}{\varepsilon^n} \int_{B(x;\varepsilon)} \omega\left(\frac{x-y}{\varepsilon}\right) f(y) dy. \end{aligned}$$

注意, 除了 $\Omega = \mathbb{R}^n$, 此时对所有的 $\varepsilon > 0$, $\Omega_\varepsilon = \Omega$, 这个特殊情况而外, $f_\varepsilon : \Omega_\varepsilon \to \mathbb{R}$ 只定义于 Ω 的真子集 Ω_ε 上.

下面的定理建立了这类正则化族的两个重要性质, 即函数 f_ε 是无限阶可微的 (正则性质); 如果 $f \in \mathcal{C}(\Omega)$, 则当 $\varepsilon \to 0$ 时, 在 Ω 的紧子集上, 函数 f_ε 一致收敛于 f (逼近性质). 在 $\Omega = \mathbb{R}^n$ 的特殊情况下, 另一个同样重要的性质将在定理 2.6-3 和 2.6-4 中给出.

定理 2.6-1 (a) 设 Ω 为 $\mathbb{R}^n$ 中的开子集, 函数 $f \in L^1_{\text{loc}}(\Omega)$, $(f_\varepsilon)_{\varepsilon>0}$ 为 f 的一个正则族. 则对任何 $\varepsilon > 0$,

$$f_\varepsilon \in \mathcal{C}^\infty(\Omega_\varepsilon),$$

而且对任何多重指标 $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n)$, $|\boldsymbol{\alpha}| = \sum_{i=1}^n \alpha_i \geqslant 1$, 和任意的 $x \in \Omega_\varepsilon$,

$$\begin{aligned} \partial^{\boldsymbol{\alpha}} f_\varepsilon(x) &= \int_\Omega \partial^{\boldsymbol{\alpha}} \omega_\varepsilon(x-y) f(y) dy \\ &= \int_{B(x;\varepsilon)} \partial^{\boldsymbol{\alpha}} \omega_\varepsilon(x-y) f(y) dy. \end{aligned}$$

(b) 再设对某个整数 $m \geqslant 1$, $f \in \mathcal{C}^m(\Omega)$. 则对 Ω 中任意给定的紧子集 K, 存在 $\varepsilon_0 = \varepsilon_0(K) > 0$, 使得 $0 < \varepsilon \leqslant \varepsilon_0$ 时 $K \subset \Omega_\varepsilon$, 对所有的 $x \in K$ 和 $0 < \varepsilon \leqslant \varepsilon_0$, $f_\varepsilon(x)$ 是唯一确定的, 且对所有的 $|\boldsymbol{\alpha}| \leqslant m$, 当 $\varepsilon \to 0$ 时

$$\sup_{x \in K} |\partial^{\boldsymbol{\alpha}} f_\varepsilon(x) - \partial^{\boldsymbol{\alpha}} f(x)| \to 0.$$

证明 (i) 在这一部分中, 固定 $\varepsilon > 0$. 设 $x \in \Omega_\varepsilon$, 对于 $1 \leqslant i \leqslant n$, e_i 为 $\mathbb{R}^n$ 的典范基中的一个向量. 因为 Ω_ε 是开集, 故而存在 $h_0 > 0$, 使得当 $|h| \leqslant h_0$ 时 $x + he_i \in \Omega_\varepsilon$. 于是, 对 $|h| \leqslant h_0$,

$$\frac{1}{h}\{f_\varepsilon(x+he_i) - f_\varepsilon(x)\} = \frac{1}{\varepsilon^n} \int_\Omega \frac{1}{h} \left\{ \omega\left(\frac{x+he_i-y}{\varepsilon}\right) - \omega\left(\frac{x-y}{\varepsilon}\right) \right\} f(y) dy.$$

由假设 $\omega \in \mathcal{C}^\infty(\mathbb{R}^n)$, 又集合 $\{(\frac{x+he_i-y}{\varepsilon}) \in \mathbb{R}^n; |h| \leqslant h_0\}$ 是紧集, 因此存在常数 M, 使得当 $|h| \leqslant h_0$ 时

$$\left| \frac{1}{h} \left\{ \omega\left(\frac{x+he_i-y}{\varepsilon}\right) - \omega\left(\frac{x-y}{\varepsilon}\right) \right\} - \frac{1}{\varepsilon} \partial_i \omega\left(\frac{x-y}{\varepsilon}\right) \right| \leqslant \frac{h}{2\varepsilon^2} M.$$

注意 $\partial_i\omega\left(\frac{x-y}{\varepsilon}\right)=\varepsilon^{n+1}\partial_i\omega_\varepsilon(x-y)$, 于是可知当 $|h|\leqslant h_0$ 时

$$\left|\frac{1}{\varepsilon^n}\int_\Omega\frac{1}{h}\left\{\omega\left(\frac{x+he_i-y}{\varepsilon}\right)-\omega\left(\frac{x-y}{\varepsilon}\right)\right\}f(y)dy-\int_\Omega\partial_i\omega_\varepsilon(x-y)f(y)dy\right|$$
$$\leqslant\frac{hM}{2\varepsilon^2}\int_{B(x;\varepsilon)}|f(y)|dy.$$

令 $h\to 0$, 即知偏导数 $\partial_i f_\varepsilon(x)$ 存在, 且对每个 $x\in\Omega_\varepsilon$,

$$\partial_i f_\varepsilon(x)=\int_\Omega\partial_i\omega_\varepsilon(x-y)f(y)dy=\int_{B(x;\varepsilon)}\partial_i\omega_\varepsilon(x-y)f(y)dy.$$

类似的论证也显然适用于偏导数 $\partial^{\boldsymbol{\alpha}}f_\varepsilon(x),|\boldsymbol{\alpha}|\geqslant 2$, 的情况.

(ii) 设 $\Omega\neq\mathbb{R}^n, f\in\mathcal{C}(\Omega)$. 给定 Ω 的任一紧子集 K, 记

$$K_0:=\{x\in\Omega;\mathrm{dist}\,(x,K)\leqslant\delta\},$$

其中 $2\delta:=\inf_{x\in K}\mathrm{dist}\,(x;\mathbb{R}^n-\Omega)>0$, 则 K_0 也是 Ω 的紧子集. 因为 $\cup_{\varepsilon>0}\Omega_\varepsilon=\Omega$ 构成 K_0 的开覆盖, 且当 $\varepsilon<\varepsilon'$ 时 $\Omega_{\varepsilon'}\subset\Omega_\varepsilon$, 因此, 存在 $\varepsilon_0=\varepsilon_0(K)>0$, 使得当任何 $\varepsilon\leqslant\varepsilon_0$ 时 $K_0\subset\Omega_\varepsilon$. 这样, 对所有的 $x\in K$ 和 $0<\varepsilon\leqslant\varepsilon_0$, $f_\varepsilon(x)=\int_{B(x;\varepsilon)}\omega_\varepsilon(x-y)f(y)dy$ 有定义. 注意到对所有的 $z\in\mathbb{R}^n$, 有 $\omega_\varepsilon(z)\geqslant 0,\int_{B(0;\varepsilon)}\omega_\varepsilon(z)dz=1$, 所以对所有的 $x\in K$ 和 $\varepsilon\leqslant\varepsilon_0$, 有

$$|f_\varepsilon-f(x)|=\left|\int_{B(0;\varepsilon)}\omega_\varepsilon(z)(f(x-z)-f(x))dz\right|$$
$$\leqslant\sup_{x\in K,z\in B(0;\varepsilon)}|f(x-z)-f(x)|.$$

于是, 由函数 f 在紧集 K_0 上的一致连续性可知当 $\varepsilon\to 0$ 时

$$\sup_{x\in K}|f_\varepsilon(x)-f(x)|\to 0.$$

(iii) 再设 $\Omega\neq\mathbb{R}^n$, 且对某个 $m\geqslant 1, f\in\mathcal{C}^m(\Omega)$. 于是, 对所有的 $x\in K$ 和所有的 $0<\varepsilon\leqslant\varepsilon_0$,

$$\partial^{\boldsymbol{\alpha}}f_\varepsilon(x)=\int_{B(x;\varepsilon)}\partial_x^{\boldsymbol{\alpha}}\omega_\varepsilon(x-y)f(y)dy=(-1)^{|\boldsymbol{\alpha}|}\int_{B(x;\varepsilon)}\partial_y^{\boldsymbol{\alpha}}\omega_\varepsilon(x-y)f(y)dy,$$

其中 $\partial_x^{\boldsymbol{\alpha}}$ 和 $\partial_y^{\boldsymbol{\alpha}}$ 分别表示关于变量 x 和 y 的偏导数. 连续做 m 次分部积分 (因为 $\mathrm{supp}\,\omega_\varepsilon(x-\cdot)\subset\overline{B(x;\varepsilon)}\subset\Omega$, 故而不要求关于 $\partial\Omega$ 的任何正则性), 即得

$$\int_{B(x;\varepsilon)}\partial_y^{\boldsymbol{\alpha}}\omega_\varepsilon(x-y)f(y)dy=(-1)^{|\boldsymbol{\alpha}|}\int_{B(x;\varepsilon)}\omega_\varepsilon(x-y)\partial^{\boldsymbol{\alpha}}f(y)dy.$$

因此,

$$|\partial^{\boldsymbol{\alpha}}f_\varepsilon(x)-\partial^{\boldsymbol{\alpha}}f(x)|=\left|\int_{B(0;\varepsilon)}\omega_\varepsilon(z)(\partial^{\boldsymbol{\alpha}}f(x-z)-\partial^{\boldsymbol{\alpha}}f(x))dz\right|$$
$$\leqslant\sup_{\substack{x\in K\\ z\in B(0;\varepsilon)}}|\partial^{\boldsymbol{\alpha}}f(x-z)-\partial^{\boldsymbol{\alpha}}f(z)|,$$

从而由 $\partial^\alpha f$ 在 K 上的一致连续性即得结论.

(iv) 如果 $\Omega = \mathbb{R}^n$, 则因此时对所有的 $x \in \mathbb{R}^n$ 和 $\varepsilon > 0, f_\varepsilon(x)$ 是唯一确定的, 只要对任意选取的 $\delta > 0$ 和 $\varepsilon_0 > 0$, 做同样的论证即可. □

注 从由积分定义的函数可微性的一般准则 (这个准则将在后面的定理 7.4-1 中建立), 即由 (i) 可得对每个 $x \in \Omega_\varepsilon$, 公式

$$\partial_i f_\varepsilon = \int_\Omega \partial_i \omega_\varepsilon(x-y) f(y) dy$$

成立.

后面将会看到, $\mathcal{C}^\infty(\Omega)$ 的子空间

$$\mathcal{D}(\Omega) := \{f \in \mathcal{C}^\infty(\Omega);\ \operatorname{supp} f \text{ 是 } \Omega \text{ 的紧子集}\}$$

将在例如 Sobolev 空间 (第 6 章) 中建立弱导数或分布等定义时起重要的作用; 注意空间 $\mathcal{D}(\Omega)$ 包含非零函数 (如以下证明中记作 $\widetilde{g}_\varepsilon$ 的函数). 但这里我们只来证明这个空间的一个十分重要的性质, 它可视为定理 2.5-3 的一个推广.

注 在文献中, 空间 $\mathcal{D}(\Omega)$ 有时被记作 $\mathcal{C}_\mathcal{C}^\infty(\Omega)$. 字母 “$\mathcal{D}$” 意味着其元素在 Ω 上分布 (6.3 节) 的定义中起关键的作用.

定理 2.6-2 设 Ω 为 $\mathbb{R}^n$ 中的开子集, 则对每个 $1 \leqslant p < \infty$, 空间 $\mathcal{D}(\Omega)$ 在 $L^p(\Omega)$ 中稠密.

证明 假设函数 $f \in L^p(\Omega)$, 给定 $\eta > 0$, 由定理 2.5-3, 存在函数 $g = g(f, \eta) \in \mathcal{C}_\mathcal{C}(\Omega)$, 使得

$$\|f - g\|_{L^p(\Omega)} \leqslant \frac{\eta}{2}.$$

设 $(g_\varepsilon)_{\varepsilon > 0}$ 为 g 的正则族. 因为 $\operatorname{supp} g$ 是 Ω 的紧子集, 重复定理 2.6-1 证明中第 (ii) 部分的证明, 可知存在 Ω 的紧子集 K_0 和 $\varepsilon_1 > 0$, 使得对任何 $\varepsilon \leqslant \varepsilon_1$,

$$\operatorname{supp} g \subset \operatorname{supp} g_\varepsilon \subset K_0 \subset \Omega_\varepsilon.$$

又因 K_0 是 Ω 的紧子集, 可知 (定理 2.6-1(b)) 当 $\varepsilon \to 0$ 时

$$\sup_{x \in K_0} |g_\varepsilon - g(x)| \to 0.$$

因此, 对充分小的 $\varepsilon > 0$,

$$\begin{aligned}
\|g - g_\varepsilon\|_{L^p(\Omega_\varepsilon)} &= \left(\int_{K_0} |g_\varepsilon(x) - g(x)|^p dx\right)^{\frac{1}{p}} \\
&\leqslant \left(\int_{K_0} dx\right)^{\frac{1}{p}} \sup_{x \in K_0} |g_\varepsilon(x) - g(x)| \leqslant \frac{\eta}{2}.
\end{aligned}$$

设 $\widetilde{g_\varepsilon}$ 是将 g_ε 在 $\Omega-\Omega_\varepsilon$ 上延拓为 0 得到的函数. 则因 $g_\varepsilon\in\mathcal{D}(\Omega_\varepsilon)$, 所以 $\widetilde{g_\varepsilon}\in\mathcal{D}(\Omega)$, 且对充分小的 $\varepsilon>0$,

$$\begin{aligned}\|f-\widetilde{g_\varepsilon}\|_{L^p(\Omega)}&=\|f-g_\varepsilon\|_{L^p(\Omega_\varepsilon)}\\&\leqslant\|f-g\|_{L^p(\Omega)}+\|g-g_\varepsilon\|_{L^p(\Omega_\varepsilon)}\leqslant\eta.\end{aligned}$$

由 $\eta>0$ 的任意性, 即得定理结论. □

注意, 定理 2.6-2 对于 $p=\infty$ 并不成立 (习题 2.6-3). 再应注意这个定理给出 Lebesgue 空间 $L^p(\Omega), 1\leqslant p<\infty$, 的另一种定义方法, 即可以把这个空间作为空间 $\mathcal{D}(\Omega)$ 关于范数 $\|\cdot\|_{L^p(\Omega)}$ 的完备化空间 (这个定义中可使用 Riemann 积分).

在本节的以下部分, 设 $\Omega=\mathbb{R}^n$, 此时, 对于给定的函数 $f\in L^1_{\text{loc}}(\mathbb{R}^n)$, f 的正则族 $(f_\varepsilon)_{\varepsilon>0}$ 中的函数 f_ε 也定义于 $\mathbb{R}^n$ 上 (当 $\Omega=\mathbb{R}^n$ 时, 对任何 $\varepsilon>0,\Omega_\varepsilon=\mathbb{R}^n$), 更特别地, 此时对任何 $x\in\mathbb{R}^n$,

$$\begin{aligned}f_\varepsilon(x)&=\int_{\mathbb{R}^n}\omega_\varepsilon(x-y)f(y)dy\\&=\int_{\mathbb{R}^n}\omega_\varepsilon(y)f(x-y)dy\end{aligned}$$

(由定理 1.16-1, $\mathbb{R}^n$ 上的上述两个积分相等).

注　函数 f_ε 实际上即函数 ω_ε 与 f 的卷积. 关于这一重要概念的详情, 见习题 2.6-4.

下面的结果建立了函数 $f\in L^p(\mathbb{R}^n), 1\leqslant p<\infty$, 的任意正则族的一个基本性质. 注意, 它也提供了当 $\Omega=\mathbb{R}^n$ 时定理 2.6-2 的另一个构造性的证明.

定理 2.6-3 (当 $1\leqslant p<\infty$ 时 $L^p(\mathbb{R}^n)$ 的正则化与逼近)　设给定函数 $f\in L^p(\mathbb{R}^n)$, $1\leqslant p<\infty, (f_\varepsilon)_{\varepsilon>0}$ 为 f 的一个正则化族, 则对任何 $\varepsilon>0$,

$$f_\varepsilon\in\mathcal{C}^\infty(\mathbb{R}^n)\cap L^p(\mathbb{R}^n),$$

而且当 $\varepsilon\to 0$ 时

$$\|f_\varepsilon-f\|_{L^p(\mathbb{R}^n)}\to 0.$$

证明　(i) 先证对任何 $\varepsilon>0$, 有 $f_\varepsilon\in L^p(\mathbb{R}^n)$ (由定理 2.6-1 已知 $f_\varepsilon\in\mathcal{C}^\infty(\mathbb{R}^n)$), 而且

$$\|f_\varepsilon\|_{L^p(\mathbb{R}^n)}\leqslant\|f\|_{L^p(\mathbb{R}^n)}.$$

如果 $p=1$, 利用 Fubini 定理 (定理 1.15-5(b)), 结合对所有的 $x\in\mathbb{R}^n,\omega_\varepsilon(x)\geqslant 0$ 和 $\int_{\mathbb{R}^n}\omega_\varepsilon(x)dx=1$, 可得

$$\begin{aligned}\int_{\mathbb{R}^n}|f_\varepsilon(x)|dx&\leqslant\int_{\mathbb{R}^n}\left(\int_{\mathbb{R}^n}\omega_\varepsilon(x-y)|f(y)|dy\right)dx\\&=\int_{\mathbb{R}^n}|f(y)|\left(\int_{\mathbb{R}^n}\omega_\varepsilon(x-y)dx\right)dy=\|f\|_{L^1(\mathbb{R}^n)}.\end{aligned}$$

如果 $1 < p < \infty$, 取 q 满足 $\frac{1}{p} + \frac{1}{q} = 1$. 由 Hölder 不等式 (定理 2.5-1) 可得对任何 $x \in \mathbb{R}$,

$$\begin{aligned}|f_\varepsilon(x)| &\leqslant \int_{B(x;\varepsilon)} \omega_\varepsilon(x-y)|f(y)|dy \\ &\leqslant \left(\int_{B(x;\varepsilon)} \omega_\varepsilon(x-y)dy\right)^{\frac{1}{q}} \left(\int_{B(x;\varepsilon)} \omega_\varepsilon(x-y)|f(y)|^p dy\right)^{\frac{1}{p}} \\ &= \left(\int_{\mathbb{R}^n} \omega_\varepsilon(x-y)|f(y)|^p dy\right)^{\frac{1}{p}}.\end{aligned}$$

因而, 再次应用 Fubini 定理, 得到

$$\begin{aligned}\int_{\mathbb{R}^n} |f_\varepsilon(x)|^p dx &\leqslant \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \omega_\varepsilon(x-y)|f(y)|^p dy\right) dx \\ &= \int_{\mathbb{R}^n} |f(y)|^p \left(\int_{\mathbb{R}^n} \omega_\varepsilon(x-y)dx\right) dy = (\|f\|_{L^p(\mathbb{R}^n)})^p.\end{aligned}$$

(ii) 其次, 再证当 $\varepsilon \to 0$ 时 $\|f_\varepsilon - f\|_{L^p(\mathbb{R}^n)} \to 0$.

假设给定 $\eta > 0$. 由定理 2.5-3, 存在函数 $g = g(f, \eta) \in \mathcal{C}_C(\Omega)$, 使得

$$\|f - g\|_{L^p(\mathbb{R}^n)} \leqslant \frac{\eta}{3}.$$

因为 supp g 是 Ω 的紧子集, 故而存在 $\mathbb{R}^n$ 的紧子集 K_0 和 $\varepsilon_1 > 0$, 使得当 $\varepsilon \leqslant \varepsilon_1$ 时

$$\operatorname{supp} g \subset \operatorname{supp} g_\varepsilon \subset K_0.$$

因此

$$\begin{aligned}\|g_\varepsilon - g\|_{L^p(\mathbb{R}^n)} &= \|g_\varepsilon - g\|_{L^p(K_0)} \\ &\leqslant \left(\int_{K_0} dx\right)^{\frac{1}{p}} \sup_{x \in K_0} |g_\varepsilon(x) - g(x)|.\end{aligned}$$

另一方面, 由定理 2.6-1(b), 当 $\varepsilon \to 0$ 时

$$\sup_{x \in K_0} |g_\varepsilon(x) - g(x)| \to 0.$$

因此, 存在 $\varepsilon_0 > 0$, 使得当 $\varepsilon \leqslant \varepsilon_0$ 时

$$\|g_\varepsilon - g\|_{L^p(\mathbb{R}^n)} \leqslant \frac{\eta}{3}.$$

注意 $f_\varepsilon - g_\varepsilon = (f-g)_\varepsilon$, 又由 (i), $\|(f-g)_\varepsilon\|_{L^p(\mathbb{R}^n)} \leqslant \|f-g\|_{L^p(\mathbb{R}^n)}$, 最后可得对任何 $\varepsilon \leqslant \varepsilon_0$,

$$\|f_\varepsilon - f\|_{L^p(\mathbb{R}^n)} \leqslant \|(f-g)_\varepsilon\|_{L^p(\mathbb{R}^n)} + \|g_\varepsilon - g\|_{L^p(\mathbb{R}^n)} + \|g - f\|_{L^p(\mathbb{R}^n)} \leqslant \eta.$$

由 $\eta > 0$ 的任意性即得定理的结论. □

下面的结果是对定理 2.6-1 和 2.6-3 的很有用的补充.

定理 2.6-4 设 Ω 是 $\mathbb{R}^n$ 的开子集, 给定函数 $f \in L^p_{\text{loc}}(\Omega), 1 \leqslant p < \infty$, 和 f 的正则族 $(f_\varepsilon)_{\varepsilon>0}$, 则对 Ω 中任何开子集 U, 如果 $\overline{U}$ 是 Ω 的紧子集, 则存在 $\varepsilon_0 = \varepsilon_0(U) > 0$, 使得 $0 < \varepsilon < \varepsilon_0$ 时 $\overline{U} \subset \Omega_\varepsilon$, 且当 $\varepsilon \to 0$ 时

$$\|f_\varepsilon - f\|_{L^p(U)} \to 0.$$

证明 设 V 是 Ω 的开子集, 满足

$$\overline{U} \subset V \text{ 且 } \overline{V} \text{ 是 } \Omega \text{ 的紧子集.}$$

由假设, $f|_V \in L^p(V)$, 且在 $\mathbb{R}^n$ 上由 $\tilde{f}|_V = f|_V$, $\tilde{f}|_{\mathbb{R}^n - V} = 0$ 定义的函数 $\tilde{f}$ 属于 $L^p(\mathbb{R}^n)$. 另外; 存在 $\varepsilon_0 > 0$, 使得当 $\varepsilon \leqslant \varepsilon_0$ 时正则族 $(f_\varepsilon)_{\varepsilon>0}$ 与 $(\tilde{f}_\varepsilon)_{\varepsilon>0}$ 中的元在 U 上一致. 于是, 当 $\varepsilon \leqslant \varepsilon_0$ 时

$$\|f_\varepsilon - f\|_{L^p(U)} = \|\tilde{f}_\varepsilon - \tilde{f}\|_{L^p(U)} \leqslant \|\tilde{f}_\varepsilon - \tilde{f}\|_{L^p(\mathbb{R}^n)}.$$

由定理 2.6-3 可得当 $\varepsilon \to 0$ 时, $\|f_\varepsilon - f\|_{L^p(U)} \to 0$. □

习题

2.6-1 称函数 $f_n \in L^1_{\text{loc}}(\Omega)$ 的序列 $(f_n)_{n=1}^\infty$ 在 $L^1_{\text{loc}}(\Omega)$ 中收敛于函数 $f \in L^1_{\text{loc}}(\Omega)$, 是指对 Ω 的任何紧子集 K, 均有

$$\lim_{n\to\infty} \|f_n - f\|_{L^1(K)} = 0.$$

证明: 存在向量空间 $L^1_{\text{loc}}(\Omega)$ 上的距离 d, 使得在 $L^1_{\text{loc}}(\Omega)$ 中 $(f_n)_{n=1}^\infty$ 收敛于 f, 当且仅当

$$\lim_{n\to\infty} d(f_n, f) = 0.$$

这个性质表明上述收敛概念定义一个距离拓扑, 即称为与半范数族 $(\|\cdot\|_{L^1(K)})_{K\in\mathcal{K}}$ 相应的 Fréchet 拓扑, 这里 $\mathcal{K}$ 为 Ω 的所有紧子集组成的集族.

提示: 类比习题 2.3-2.

2.6-2 (1) 函数 $\theta : \mathbb{R}^n \to \mathbb{R}$ 定义为当 $|x| < 1$ 时 $\theta(x) = e^{\frac{1}{|x|^2-1}}$, 当 $|x| \geqslant 1$ 时 $\theta(x) = 0$, 证明 θ 在 $\mathbb{R}^n$ 中无限阶可微.

(2) 对于函数 θ, 在点 $x_0 \in \mathbb{R}^n$, $|x_0| = 1$, 处的带有积分余项的 Taylor 公式是什么形式?

2.6-3 证明子空间 $\mathcal{D}(\Omega)$ 在 $L^\infty(\Omega)$ 中并不稠密.

2.6-4 设给定两个函数 $f \in L^1(\mathbb{R}^n)$ 和 $g \in L^p(\mathbb{R}^n), 1 \leqslant p \leqslant \infty$.

(1) 证明对几乎所有的 $x \in \mathbb{R}^n$, 函数 $y \in \mathbb{R}^n \to f(x-y)g(y)$ 在 $\mathbb{R}^n$ 上可积, 因而对几乎所有的 $x \in \mathbb{R}^n$ 定义了一个实数

$$(f * g)(x) := \int_{\mathbb{R}^n} f(x-y)g(y)dy.$$

这样便定义了函数 $f * g : \mathbb{R}^n \to \mathbb{R}$, 称之为 f 和 g 的卷积.

(2) 证明

$$f*g \in L^p(\mathbb{R}^n) \text{ 且 } \|f*g\|_{L^p(\mathbb{R}^n)} \leqslant \|f\|_{L^1(\mathbb{R}^n)}\|g\|_{L^p(\mathbb{R}^n)}.$$

注 由此, 双线性映射

$$(f,g) \in L^1(\mathbb{R}^n) \times L^p(\mathbb{R}^n) \to f*g \in L^p(\mathbb{R}^n)$$

是连续的 (定理 2.11-1).

2.6-5 (1) 设 $\varphi \in L^\infty(\mathbb{R}^n)$ 为具有紧支集且几乎处处非负的函数, 证明: 存在 $\mathbb{R}^n$ 的有界开子集 U 和函数 $\varphi_k : U \to \mathbb{R}$, 其中 $k \geqslant 1$, 满足以下性质:

$$\begin{aligned}
&\varphi_k \in \mathcal{D}(U),\ \text{在 } U \text{ 上 } \varphi_k \geqslant 0,\\
&\text{对任何 } k \geqslant 1, \|\varphi_k\|_{L^\infty(U)} \leqslant \|\varphi\|_{L^\infty(U)},\\
&\text{对几乎所有的 } x \in U,\ \text{当 } k \to \infty \text{ 时 } \varphi_k(x) \to \varphi(x).
\end{aligned}$$

(2) 设 Ω 为 $\mathbb{R}^n$ 的一个开子集, 函数 $f \in L^1_{\text{loc}}(\Omega)$, 使得对一切在 Ω 上非负的函数 $\varphi \in \mathcal{D}(\Omega)$ 均有

$$\int_\Omega f\varphi dx \geqslant 0,$$

证明: 在 Ω 上几乎处处有 $f \geqslant 0$.

提示: 首先, 利用 (1) 证明对于在 Ω 上几乎处处非负且具有紧支集的任何 $\varphi \in L^\infty(\Omega)$ 均有 $\int_\Omega f\varphi dx \geqslant 0$, 然后, 再证明对 Ω 上任意给定的开子集 V, 若 $\overline{V}$ 为紧集, 则有 $dx\text{-}meas\{x \in V; f(x) < 0\} = 0$.

2.7 紧性和有限维赋范向量空间; F. Riesz 定理

本节的目的是回顾一下有限维赋范向量空间的基本性质, 其中大部分与紧性概念有关.

作为开始, 下一个定理中的性质 (a) 本质上说明有限维向量空间只有一个范数拓扑, 因而它可以由范数 $\|\cdot\|_p$, $1 \leqslant p \leqslant \infty$, 定义 (定理 2.2-2). 性质 (b) 和 (c) 将有限维空间上紧集的特征从范数 $\|\cdot\|_p$, $1 \leqslant p \leqslant \infty$, 之一的情况 (定理 1.13-5 和 2.2-2(b)) 推广到任意范数的情况. 性质 (d) 是有限维子空间的一个重要的拓扑性质. 值得注意的是性质 (b)(c) 和 (d) 的证明均依赖性质 (a).

定理 2.7-1 (a) 有限维向量空间 X 上任意两个范数 $\|\cdot\|$ 和 $\|\cdot\|'$ 均是等价的, 即由 $\|\cdot\|$ 和 $\|\cdot\|'$ 在 X 上导出的拓扑是一致的.

(b) 任何有限维赋范向量空间都是可分的.

(c) 有限维赋范向量空间中的子集为紧集的充要条件是它为有界闭集.

(d) 赋范向量空间 X 的有限维子空间在 X 中闭.

证明 (i) 设 $(e_i)_{i=1}^n$ 为 X 的基. 显然, 只要证明 X 上的任意范数均等价于特殊的范数 $\|\cdot\|_1 : x = \sum_{i=1}^n x_i e_i \to \sum_{i=1}^\infty |x_i|$ (定理 2.2-4). 为此, 首先注意到对任何 $x \in X$,

$$\|x\| = \left\| \sum_{i=1}^{\infty} x_i e_i \right\| \leqslant C_1 \|x\|_1,$$

其中 $C_1 := \max_{1 \leqslant i \leqslant n} \|e_i\|$.

再考察函数

$$f : x \in (X, \|\cdot\|_1) \to f(x) := \|x\| \in \mathbb{R},$$

令

$$K := \{y \in X; \|y\|_1 = 1\}.$$

因为对任何 $x, y \in X$,

$$|f(x) - f(y)| = |\,\|x\| - \|y\|\,| \leqslant \|x - y\| \leqslant C_1 \|x - y\|_1,$$

所以 f 是 X 上的连续函数. 又 K 作为距离空间 (X, d_1) 的有界闭子集, 它是 X 中的紧集 (定理 1.13-5). 因此, 存在 $y_0 \in K$, 使得 $f(y_0) = \inf_{y \in K} f(y)$ (定理 1.13-6), 又因 $y_0 \neq 0$, 故而 $\frac{1}{C} := f(y_0) = \|y_0\| > 0$, 因此, 由 $\|y\|_1 = 1$ 得到

$$\|y\| \geqslant \frac{1}{C}.$$

给定任意的非零向量 $x \in X, y := \frac{x}{\|x\|_1}$ 满足 $\|y\|_1 = 1$, 所以, $\|y\| = \frac{\|x\|}{\|x\|_1} \geqslant \frac{1}{C}$. 于是可得对任何 $x \in X$,

$$\|x\|_1 \leqslant C\|x\|.$$

这样, 由定理 2.2-4, 由 $\|\cdot\|$ 和 $\|\cdot\|'$ 导出的拓扑是一致的. 这就证明了 (a), 再结合定理 2.2-2(b), 即证得 (b).

(ii) 设 K 为有限维赋范向量空间 $(X, \|\cdot\|)$ 中的有界闭子集. 由 (a), K 在 $(X, \|\cdot\|)$ 中是有界闭的, 从而由定理 1.13-5, 它是 $(X, \|\cdot\|_1)$ 中的紧集. 又由 (a), $(X, \|\cdot\|)$ 与 $(X, \|\cdot\|_1)$ 上的拓扑相同, 所以 K 是 $(X, \|\cdot\|)$ 中的紧集. (c) 的必要性部分在任何距离空间中成立 (定理 1.13-1), 因此在 $(X, \|\cdot\|)$ 中成立 (不论 X 是否为有限维空间). (c) 得证.

(iii) 设 Y 是赋范向量空间 $(X, \|\cdot\|)$ 的有限维子空间, $(e_i)_{i=1}^n$ 是 Y 的基, $(y^k)_{k=1}^\infty$ 为在 X 中收敛的向量 $y^k = \sum_{i=1}^n y_i^k e_i \in Y$ 的序列.

由 (a), 存在常数 C, 使得对任何 $k, \ell \geqslant 1$,

$$\sum_{i=1}^{n} |y_i^k - y_i^\ell| = \|y^k - y^\ell\|_1 \leqslant C\|y^k - y^\ell\|,$$

由 $(y^k)_{k=1}^{\infty}$ 的收敛性可知它是 Cauchy 序列 (定理 1.12-1(b)), 因此, 对数量 $y_i \in \mathbb{K}$, $1 \leqslant i \leqslant n$, 每个序列 $(y_i^k)_{k=1}^{\infty}$ 均为 Cauchy 序列. 因为数域 $\mathbb{K}$ 是完备的, 故而存在 $y_i \in \mathbb{K}$, $1 \leqslant i \leqslant n$, 使得当 $k \to \infty$ 时 $y_i^k \to y_i$, 这样 $\lim_{k\to\infty} \|y^k - y\|_1 = 0$, 其中 $y := \sum_{i=1}^n y_i e_i$.

由 (a), 存在常数 C_1, 使得 $\|y^k - y\| \leqslant C_1 \|y^k - y\|_1$ 对 $k \geqslant 1$ 均成立 (向量 $y^k, k \geqslant 1$ 和 y 均属于 Y), 从而 $\lim_{k\to\infty} \|y^k - y\| = 0$, 于是, 序列 $(y^k)_{k=1}^{\infty}$ 收敛于向量 $y \in Y$, 即 Y 是闭的. 这就证得了 (d). □

作为一个应用, 设 $\mathcal{P}_n$ 表示所有次数 $\leqslant n$ 的实多项式 $p : x \in \mathbb{R} \to p(x) = \sum_{j=0}^n c_j(p) x^j$ 组成的空间. 由定理 2.7-1(a) 存在 (依赖 n 的) 常数 C 和 C_1, 使得对一切 $p \in \mathcal{P}_n$,

$$\sum_{j=0}^n |c_j(p)| \leqslant C \sup_{0 \leqslant x \leqslant 1} |p(x)|,$$

$$\sup_{0 \leqslant x \leqslant 1} |p(x)| \leqslant C_1 \sum_{j=0}^n |c_j(p)|.$$

值得注意的是后一不等式容易直接导出 (取 $C_1 = 1$), 而前一不等式则不然.

我们还将证明定理 2.7-1(a) 中建立起来的范数的等价性实际上是有限维向量空间的特征. 为此, 不用奇怪的是将要应用选择公理, 原因是 Hamel 基 (证明中要用到) 的存在性依赖这个公理.

定理 2.7-2 *设 X 为任意的无限维向量空间, 则在 X 上存在不等价的范数.*

证明 设 $(e_i)_{i \in I}$ 为 X 的 Hamel 基 (2.1 节), 即任何向量 $x \in X$ 均可唯一地表示为 $x = \sum_{j \in J(x)} x_j e_j$, 其中 $J(x)$ 是 I 的有限子集. 定义映射 $\|\cdot\|_1 : X \to \mathbb{R}$ 和 $\|\cdot\|_\infty : X \to \mathbb{R}$ 为

$$\|x\|_1 := \sum_{j \in J(x)} |x_j|, \quad \|x\|_\infty := \max_{j \in J(x)} |x_j|.$$

可以直接验证它们都是 X 上的范数. 因为 I 是无限集 (由假设, X 为无限维的), 所以 Hamel 基 $(e_i)_{i \in I}$ 包含一个可数无限的子族 $(e_j)_{j=1}^{\infty}$ (定理 1.5-3(a)). 考察序列 $(x_n)_{n=1}^{\infty}$, 其中 $x_n := \sum_{j=1}^n \frac{1}{n} e_j$, 于是 $\|x_n\|_1 = 1$, $\|x_n\|_\infty = \frac{1}{n}$, $n \geqslant 1$. 因此, 不存在常数 C, 使得对一切 x 均有 $\|x\|_1 \leqslant C \|x\|_\infty$. □

由定理 2.7-1(c), 在有限维赋范向量空间 $(X, \|\cdot\|)$ 中, 单位球面 $\{x \in X; \|x\| = 1\}$ 作为 X 中一个特殊的有界闭集, 它是一个紧集. 值得注意的是由于下述基本定理, 这个性质也是有限维向量空间的特征.

定理 2.7-3 (F. Riesz 定理) *赋范向量空间 $(X, \|\cdot\|)$ 是有限维空间的充要条件为 X 的单位球面是紧集.*

证明 设单位球面

$$K := \{x \in X; \|x\| = 1\}$$

在 $(X, \|\cdot\|)$ 中紧. 于是, 存在有限个点 $x_i \in X$, $1 \leqslant i \leqslant n$, 使得 $K \subset \cup_{i=1}^{n} B(x_i; \frac{1}{2})$ (1.13 节).

思路是证明 X 与有限维向量空间

$$Y := \text{Span}\,(x_i)_{i=1}^{n}$$

是一致的. 为此, 只要证明任意给定的 $x \in X$, 均有

$$\inf_{y \in Y} \|x - y\| = 0,$$

由此即得 $x \in \overline{Y}$, 又由于 Y 是有限维空间, 故而 $\overline{Y} = Y$ (定理 2.7-1(d)).

对给定的 $x \in X$, 如果 $x \in Y$, 则已证得结论. 否则, 任取 $y \in Y$, 令 $\tilde{x} := \frac{x}{\|x-y\|}, \tilde{y} := \frac{y}{\|x-y\|}$. 因为 $\tilde{x} - \tilde{y} \in K$, 故存在 $i_0, 1 \leqslant i_0 \leqslant n$, 使得 $\tilde{x} - \tilde{y} \in B(x_{i_0}; 1/2)$, 所以

$$\|x - \|x - y\|(\tilde{y} + x_{i_0})\| = \|x - y\|(\|(\tilde{x} - \tilde{y}) - x_{i_0}\|) < \frac{1}{2}\|x - y\|.$$

因为 $\tilde{y}$ 和 x_{i_0} 属于 Y, 所以向量 $y_1 := \|x - y\|(\tilde{y} + x_{i_0})$ 属于 Y.

综上可得, 如果 $x \notin Y$, 则对给定的任何 $y \in Y$, 必定存在 $y_1 \in Y$, 使得 $\|x - y_1\| < \frac{1}{2}\|x - y\|$. 由递推, 即知存在向量 $y_n \in Y$, 使得

$$\|x - y_n\| < \frac{1}{2^n}\|x - y\|.$$

因此, $\inf_{y \in Y} \|x - y\| = 0$, 这就证得定理的充分性部分 (显然证明中的 $\frac{1}{2}$ 可以用开区间 $]0, 1[$ 中的任何数取代).

定理的必要性部分已在定理 2.7-1(c) 中证得. □

注意, F. Riesz 定理也可等价地叙述为: 赋范向量空间是有限维空间当且仅当闭单位球是紧集 (因为此时单位球面作为闭单位球的闭子集, 它也是紧集).

习题

2.7-1 设 Y 是赋范向量空间 $(X, \|\cdot\|)$ 的有限维子空间.

(1) 证明: 对任意给定的 $x \in X$, 存在 (未必唯一) 向量 $\tilde{y} \in Y$, 使得

$$\|x - \tilde{y}\| = \inf_{y \in Y} \|x - y\|.$$

(2) 又设 $(X, \|\cdot\|)$ 是严格凸的, 即当 $\|x\| = \|y\| = 1, x \neq y$ 时, $\left\|\frac{x+y}{2}\right\| < 1$, 证明 (1) 中的向量 $\tilde{y} \in Y$ 是唯一的.

(3) 证明: 空间 $(\mathbb{K}^n, \|\cdot\|_p)$ (2.2 节) 对任何 $1 < p < \infty$ 是严格凸的, 当 $p = 1$ 和 $p = \infty$ 时则不然.

2.7-2 证明无限维赋范向量空间的任何紧子集的内部是空集.

2.7-3 以下, 在空间 $\mathcal{C}[0, 2\pi]$ 上赋以 sup 范数. 设函数 $g_n \in \mathcal{C}[0, 2\pi], n \geqslant 1$, 定义为 $g_n(\theta) := \sin n\theta$, $0 \leqslant \theta \leqslant 2\pi$, 直接证明, 即不借助于 F. Riesz 定理, 序列 $(g_n)_{n=1}^{\infty}$ (显然是有界的) 不包含任何收敛子列.

2.8 有限维赋范向量空间中紧性的应用; 代数学基本定理

代数学基本定理叙述为任何次数 $n \geqslant 1$ 的实或复的多项式, 即实系数或复系数多项式, 至少有一个复根. 由公式 $z^k - z_0^k = (z - z_0)(z^{k-1} + z^{k-2}z_0 + \cdots + z_0^{k-1})$ 可知这样的多项式按重数计恰有 n 个复根. 对这个当初难以证明的结果的探索, 长期以来吸引了数学家们的强烈关注.

古希腊人已经掌握了二次实多项式实数根 (如果存在) 的计算公式. 然而直到 1545 年, Girolamo Cardano 才发表了实际上由 Scipione del Ferro 和 Nicolo Tartaglia 获得的三次多项式的求根公式. 其实在 1540 年 Lodovico Ferrari 已经建立了四次多项式的求根公式 (但直到很久以后才发表). 然而这些公式[11] 却被他们尚未达到但又是所要求的对复数理论的透彻理解所掩盖, 仍然处于未成熟的阶段.

Niels Henrick Abel[12] 于 1823 年, 在不可置信的 21 岁时, 证明了对一般的五次多项式不存在这样的求根公式 (所谓公式即只包含初等运算和开 p 次方, $p \geqslant 2$, 的有限表达式), 接着由 Evariste Galois 在 1832 年完成最后一击. 他在同样不可置信的 20 岁时, 证明了对任何次数 $\geqslant 5$ 的一般多项式不存在这样的求根公式. Galois 的发现[13] 是数学史上最伟大的成就之一.

与此同时, 人们还多方尝试并不寻求特殊的公式, 而借助于分析的整体力量来建立代数学基本定理. 许多数学家致力于这个方向的探索, 这里同样出现许多需要熟练地使用神秘的复数的困难. 这些数学家中, Jean Le Rond d'Alembert 被公认为在证明代数学基本定理方面于 1746 年发表了第一个 “严格的尝试”, 尽管在他的证明中尚有

[11] 寻求满足 $p(z) = 0$ 的 $z \in \mathbb{C}$, 此处 p 为 n 次多项式, 的一个聪明的途径是 (在消去 $n-1$ 次单项式后) 寻求一个 $n \times n$ 循环矩阵 (迹为 0), 其特征多项式恰为 p, 对 $n = 3$ 和 $n = 4$, 这个过程导致 p 的根的显式表示. 详细的叙述可见:

I. Kra; S. R. Simanca [2012]: On circulant matrices. Notices of the American Mathematical Society 59, 368–377.

[12] 聚焦于 Abel (1802—1829) 的有一本动人的传记:

A. Stubhaug [2000]: Niels Henrik Abel and his Times-Called Too Soon by Flames Afar, Springer, Heidelberg (由挪威文翻译而来).

[13] 似乎没有关于 Galois (1811—1832) 这位同样杰出的数学家的权威性的传记 (类似于 Stubhaug 的关于 Abel 的传记). 以现代的观点重新考察 Galois 的全部数学贡献的学术性的记载见:

P. M. Neumann [2011]: The Mathematical Writings of Évariste Galois. European Mathematical Society.

缺陷. 这也解释了为什么代数学基本定理有时被称作 d'Alembert 定理. 对于实多项式的第一个正确的证明 (按现行标准的 "正确") 是由 Carl-Friedrich Gauss 于 1816 年发表的 (其初次但并不完备的尝试是在他 1799 年博士论文中). 他又在 1849 年对复多项式给出了第一个正确的证明.

附带指出. 一般虽常称为 "代数学基本定理", 实际上却是一个分析的定理!

下面将给出的非常简单的证明[14] 依赖两个与紧性相关的基本性质, 即有限维赋范向量空间 $(\mathbb{R}^2; \|\cdot\|_2)$ 中紧子集的特征 (定理 2.7-1(c)) 和紧集上的连续函数达到其最小值 (定理 1.13-6).

定理 2.8-1 (代数学基本定理) 任何次数 $\geqslant 1$ 的复多项式在 $\mathbb{C}$ 中至少有一个根.

证明 设 $p: \mathbb{C} \to \mathbb{C}$ 为次数 $n \geqslant 1$ 的复多项式,

$$p(z) := a_n z^n + \cdots + a_1 z + a_0, \quad z \in \mathbb{C},$$

其中 $a_0, a_1, \cdots, a_n$ 均为复数, $a_n \neq 0$.

(i) 利用紧性, 先证明存在 $z_0 \in \mathbb{C}$, 使得

$$|p(z_0)| = \inf_{z \in \mathbb{C}} |p(z)|.$$

为此, 注意到

$$\lim_{|z| \to \infty} |p(z)| = \lim_{|z| \to \infty} \left(|z^n| \left| a_n + \frac{a_{n-1}}{z} + \cdots + \frac{a_0}{z^n} \right| \right) = \infty,$$

所以存在 $r > 0$, 使得当 $|z| \geqslant r$ 时

$$|p(z)| \geqslant |p(0)|.$$

显然可以将 $(\mathbb{C}, |\cdot|)$ 与 $(\mathbb{R}^2, \|\cdot\|_2)$ 视为同一拓扑空间, 即当 $z = x + iy$ 时 $|z| = \|(x, y)\|_2$. 由定理 2.7-1(c) 集合

$$K := \{z \in \mathbb{C}; |z| \leqslant r\}$$

是紧集. 又因为

$$\big| |p(z_1)| - |p(z_2)| \big| \leqslant \sum_{k=0}^{n} |a_k| \, |z_1^k - z_2^k|,$$

故而函数 $z \in \mathbb{C} \to |p(z)| \in \mathbb{R}$ 是连续的.

这样, (由定理 1.13-6) 存在 $z_0 \in K$, 使得

$$\inf_{z \in K} |p(z)| = |p(z_0)|.$$

因为当 $|z| \geqslant r$ 时

$$|p(z)| \geqslant |p(0)| \geqslant |p(z_0)|,$$

[14] 这个证明见 Schwartz [1991, 定理 2.7.10].

因此对所有的 $z \in \mathbb{C}, |p(z)| \geqslant |p(z_0)|$.

如果 $p(z_0) = 0$, 则定理已得证, 所以只要考虑 $p(z_0) \neq 0$ 的情况.

(ii) 利用复数的初等代数性质, 证明如果 $z_0 \in \mathbb{C}, p(z_0) \neq 0$, 则存在 $z_1 \in \mathbb{C}$, 使得 $|p(z_1)| < |p(z_0)|$. 由 z_0 处的 Taylor 展开式可知, 存在整数 $k, 1 \leqslant k \leqslant n$ 和复数 $c_k, c_{k+1}, \cdots, c_n$, 其中 $c_k \neq 0$, $c_n = a_n \neq 0$, 使得

$$p(z) = p(z_0) + c_k(z - z_0)^k + c_{k+1}(z - z_0)^{k+1} + \cdots + c_n(z - z_0)^n.$$

下面证明当 z 在以 z_0 为中心、半径 ε 充分小的圆周上变动时, $|p(z)|$ 严格地小于 $|p(z_0)|$. 设 $\varepsilon > 0$ 满足

$$|c_{k+1}|\varepsilon + \cdots + |c_n|\varepsilon^{n-k} < |c_k|, \quad |c_k|\varepsilon^k < |p(z_0)|.$$

其中约定当 $k = n$ 时, 上述不等式的左边为零. 当 z 在圆周 $\Gamma := \{z \in \mathbb{C}; |z - z_0| = \varepsilon\}$ 运动一周, 点 $\{p(z_0) + c_k(z - z_0)^k\}$ 在中心为 $p(z_0)$、半径为 $|c_k|\varepsilon^k$ 的圆周上运动 k 周 (图 2.8-1). 因为 $|c_k|\varepsilon^k < |p(z_0)|$, 存在 $z_1 \in \Gamma$, 使得点 $\{p(z_0) + c_k(z_1 - z_0)^k\}$ 落在联结 $\mathbb{C}$ 的原点与 $p(z_0)$ 的线段上 (解方程 $(z - z_0)^k = -\varepsilon^k \frac{|c_k| p(z_0)}{c_k |p(z_0)|}$ 即可), 因此

$$|p(z_0) + c_k(z_1 - z_0)^k| = |p(z_0)| - |c_k|\varepsilon^k.$$

这样,

$$\begin{aligned}|p(z_1)| &\leqslant |p(z_0) + c_k(z_1 - z_0)^k| + |c_{k+1}(z_1 - z_0)^{k+1} + \cdots + c_n(z_1 - z_0)^n| \\ &< |p(z_0)|.\end{aligned}$$

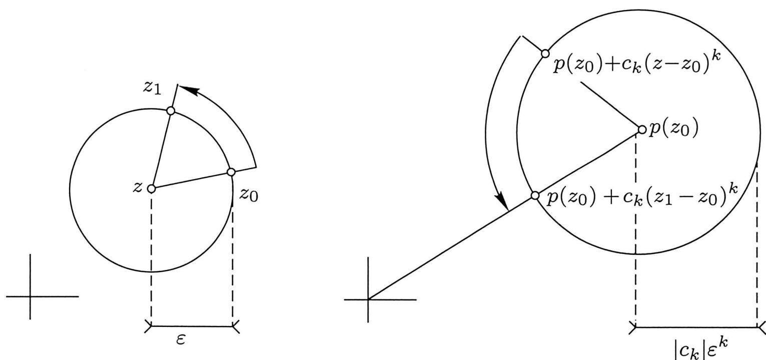


图 2.8-1　定理 2.8-1 证明中 z_1 构造的几何解释

(iii) 结合 (i) 和 (ii), 即得代数学基本定理. □

注　在许多教科书中, 上述证明中的 (ii) 被 Liouville 定理的应用所取代, 这个定理是单复变函数论的一个基本结果, 它叙述为整个复平面 $\mathbb{C}$ 上的有界解析函数必为常

数. 因此, 如果 p 是次数 $\geqslant 1$ 的多项式且在 $\mathbb{C}$ 上没有根, 则 $\frac{1}{p}$ 在 $\mathbb{C}$ 上解析, 又由 (i), 对一切 $z \in \mathbb{C}$, $\frac{1}{|p(z)|} \leqslant \frac{1}{|p(z_0)|}$, 故有界, 由 Liouville 定理, p 为常数函数, 此为矛盾.

2.9 赋范向量空间上的连续线性算子; 空间 $\mathcal{L}(X;Y), \mathcal{L}(X)$ 和 X'

下面设 X 和 Y 是同一个域 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的向量空间, 用记号 0 同时表示 X 中的零向量和 Y 中的零向量.

设映射 $A : X \to Y$ 满足对所有 $x, y \in X$ 和 $\alpha \in \mathbb{K}$,

$$A(x+y) = A(x) + A(y), \quad A(\alpha x) = \alpha A(x)$$

(因此 $A(0) = 0$), 则称 A 为由 X 到 Y 的线性算子, 当 $Y = \mathbb{K}$ 时称为线性泛函或线性型. 在不会引起误解时, 对线性算子 A, 简记 $A(x)$ 为 Ax; 简记两个线性算子 A 和 B 的复合 $A \circ B$ 为 AB; 如果当 $X = Y$ 时, 简记 $A \circ A, A \circ A \circ A$ 等为 A^2, A^3 等, 记 $A^\circ = I_X$.

如果 $\mathbb{K} = \mathbb{C}$, 映射 $A : X \to Y$ 满足对一切 $x, y \in X, \alpha \in \mathbb{K}$,

$$A(x+y) = A(x) + A(y), \quad A(\alpha x) = \bar{\alpha} A(x),$$

其中 $\bar{\alpha}$ 为 α 的复共轭, 则称 A 为半线性的. 这个概念在复内积空间中将会很自然地产生 (4.1 节).

设 $A : X \to Y$ 为线性算子. 称 X 的子集

$$\operatorname{Ker} A := \{x \in X; Ax = 0\}$$

为 A 的核, 称 X 在 A 作用下的直接像 $A(X)$ (1.2 节) 为 A 的值域, 也记作 $\operatorname{Im} A$, 即

$$\operatorname{Im} A := A(X) = \{y \in Y;\ 存在\ x \in X,\ 使得\ y = Ax\}.$$

显然, $\operatorname{Ker} A$ 是 X 的子空间, $\operatorname{Im} A$ 是 Y 的子空间.

注 尽管可能会有误解, 仍用同样的记号 Im 表示复数 z 的虚部 $\operatorname{Im} z$.

线性算子的下列初等性质会被经常用到.

定理 2.9-1 (a) 线性算子 $A : X \to Y$ 是单射当且仅当 $\operatorname{Ker} A = \{0\}$.

(b) 如果线性算子 $A : X \to Y$ 为单射, 则映射 $A : X \to \operatorname{Im} A$ 的逆映射 $B : \operatorname{Im} A \to X$ 是 $\operatorname{Im} A$ 到 X 上的线性算子.

证明 映射 $A : X \to Y$ 是单射, 当且仅当 $Ax = A\tilde{x}$ 时 $x = \tilde{x}$, 因此, 当 A 是线性算子时, 当且仅当 $Ax = 0$ 时 $x = 0$.

如果 A 是单射, 则 $BA = I_X$, 其中 I_X 是 X 上的恒等映射. 给定任意两个向量 $y,\tilde{y} \in \operatorname{Im} A$, 存在唯一确定的向量 $x,\tilde{x} \in X$, 使得 $y = Ax, \tilde{y} = A\tilde{x}$. 因此, 对任意的数 $\beta,\widetilde{\beta} \in \mathbb{K}$,

$$\begin{aligned}B(\beta y + \widetilde{\beta}\widetilde{y}) &= B(\beta Ax + \widetilde{\beta}A\tilde{x})\\ &= B(A(\beta x + \widetilde{\beta}\tilde{x})) = \beta x + \widetilde{\beta}\tilde{x} = \beta By + \widetilde{\beta}B\tilde{y}.\end{aligned}$$ □

定义线性算子的加法和数乘为

$$A + B : x \in X \to Ax + Bx \in Y,$$
$$\alpha A : x \in X \to \alpha(Ax) \in Y,$$

于是 $X \to Y$ 的所有线性算子的集合成为同一个域 $\mathbb{K}$ 上的线性空间. 其零向量为 $0 : x \in X \to 0 \in Y$ (这个零向量的记号与 X 和 Y 的零向量记号相同).

设 X 为 $\mathbb{K}$ 上的向量空间, $A : X \to X$ 为线性算子, 如果存在向量 $p \in X, p \neq 0$, 使得

$$Ap = \lambda p,$$

则称 $\lambda \in \mathbb{K}$ 为 A 的一个特征值, 称非零向量 p 为 A 相应于特征值 λ 的一个特征向量, 称 X 的子空间

$$\{p \in X; Ap = \lambda p\} \neq \{0\}$$

为相应于特征值 λ 的特征子空间.

注意, 当且仅当 0 并非 A 的特征值时, A 为单射.

当 X 和 Y 为赋范向量空间, 其上赋以范数拓扑时 (2.2 节), 连续线性算子, 或当 $Y = \mathbb{K}$ 时的连续线性泛函, 具有一些特殊的性质, 下面的定理列出最初等而基本的一些性质. 为记号简洁起见, $\|\cdot\|$ 表示未必是同一个向量空间的范数. 这种记号是可以避免任何误解的.

定理 2.9-2 设 X 和 Y 为两个赋范向量空间, $A : X \to Y$ 为线性算子, 则下列性质等价:

(a) 线性算子 A 在 X 上连续.

(b) 线性算子 A 在 X 的原点处连续.

(c) 存在常数 $C > 0$, 使得对任何 $x \in X$,

$$\|Ax\| \leqslant C\|x\|.$$

(d) X 的任何有界子集在 A 作用下的直接像是 Y 中的有界子集.

证明 显然, 由 (a) 可导出 (b).

如果 (b) 成立, Y 的闭单位球在 A 下的逆像包含一个以 X 的原点为中心的闭球; 设其半径为 $\frac{1}{C}>0$. 于是, 对任何非零向量 $x\in X$, 有 $\|A\left(\frac{x}{C\|x\|}\right)\|\leqslant 1$, 即 $\|Ax\|\leqslant C\|x\|$. 这样由 (b) 导出 (c).

如果 (c) 成立, 因为 X 的任何有界子集 B 均可包含在一个以 X 的原点为中心, 半径 $r=r(B)>0$ 的球中, 其直接像包含在一个以 Y 的原点为中心、半径为 Cr 的球中; 从而 $A(B)$ 是有界的. 这样由 (c) 导出 (d).

如果 (d) 成立, X 中闭单位球的直接像在 Y 中有界, 即存在 $M>0$, 使得当 $\|x\|\leqslant 1$ 时 $\|Ax\|\leqslant M$. 对任意给定的 $x_0\in X$ 和 $\varepsilon>0$, 取 $\delta:=\frac{\varepsilon}{M}$. 则当 $\|x-x_0\|\leqslant\delta$ 时 $\frac{1}{\delta}\|A(x-x_0)\|=\|A\left(\frac{x-x_0}{\delta}\right)\|\leqslant M$, 从而 $\|Ax-Ax_0\|\leqslant\varepsilon$, 这就证得了 (a). □

性质 (d) 说明了赋范向量空间中连续线性算子为何又被称为有界线性算子.

设 X 为 Y 的一个子空间. 记号

$$X\hookrightarrow Y$$

表示由 X 到 Y 的典则嵌入 (1.2 节). 它显然是线性的, 而且还是连续的, 或可等价地表述为 (定理 2.9-2) 存在常数 C, 使得对任何 $x\in X$,

$$\|x\|_Y\leqslant C\|x\|_X.$$

定理 2.9-3 设 X 和 Y 为两个赋范向量空间.

(a) 由 X 到 Y 的任何连续线性算子都是一致连续的.

(b) 如果 X 是有限维空间, 则由 X 到 Y 的任何线性算子均是连续的.

证明 连续线性算子 $A:X\to Y$ 的一致连续性是因为对任何 $x,\tilde{x}\in X$, 均有 $\|Ax-A\tilde{x}\|\leqslant C\|x-\tilde{x}\|$ (定理 2.9-2(c)).

再设 X 是有限维空间, $(e_i)_{i=1}^n$ 是 X 的基. 于是, 对任何向量 $x=\sum_{i=1}^n x_ie_i\in X$,

$$\|Ax\|=\|A\left(\sum_{i=1}^n x_ie_i\right)\|\leqslant C_1\|x\|_1,$$

其中 $C_1:=\max_{1\leqslant i\leqslant n}\|Ae_i\|, \|x\|_1=\sum_{i=1}^n|x_i|$ (定理 2.2-2). 由定理 2.9-2(c), 结合有限维空间中任何两个范数等价 (定理 2.7-1(a)), 即证得 A 的连续性. □

作为应用, 用 $\mathcal{P}_n$ 记所有次数 $\leqslant n$ 的实多项式组成的空间, 赋以范数 $p\in\mathcal{P}_n\to\sup_{-1\leqslant x\leqslant 1}|p(x)|$, 线性算子 $A:\mathcal{P}_n\to\mathcal{P}_n$ 定义为 $Ap=p'$. 由定理 2.9-2(c) 和 2.9-3(b) 可知存在常数 $C(n)$, 使得对任何 $p\in\mathcal{P}_n$,

$$\sup_{-1\leqslant x\leqslant 1}|p'(x)|\leqslant C(n)\sup_{-1\leqslant x\leqslant 1}|p(x)|.$$

可以进一步证明这个不等式中 "最佳" (即最小) 的常数 $C(n)$ 为 n^2 (如对应于特殊的多项式 $x\in\mathbb{R}\to\cos(n\text{Arccos }x)$ 等式成立). 这个结果即著名的 Markoff 不等

式[15], 其证明并非平凡[16]. 相应于其他范数的类似的不等式还可由定理 2.9-2(c) 导出. 例如, 对任何整数 $n \geqslant 0$ 和任意的 $r>1$, 存在常数 $C(n,r)$, 使得对一切 $p \in \mathcal{P}_n$, 成立[17]

$$\left(\int_{-1}^{1}|p'(x)|^r dx\right)^{\frac{1}{r}} \leqslant C(n,r)\left(\int_{-1}^{1}|p(x)|^r dx\right)^{\frac{1}{r}}.$$

与定理 2.9-3 性质 (b) 相反, 当空间 X 是无限维时, 线性算子 $A: X \to Y$ 的连续性和空间 Y 的维数无关.

为说明即便当 $\dim Y = 1$ 时线性算子仍然可能不连续, 如可考察任意次数的实多项式全体组成的空间 $\mathcal{P}$, 其上赋以范数 $\|\cdot\|: p \in \mathcal{P} \to \|p\| = \sup_{-1\leqslant x\leqslant 1}|p(x)|$, 令线性泛函 $f: \mathcal{P} \to \mathbb{R}$ 定义为 $f(p)=p(3)$. 对每个整数 $k \geqslant 0$, 多项式 $p_k \in \mathcal{P}$ 定义为 $p_k(x)=\left(\frac{x}{2}\right)^k, x \in \mathbb{R}$. 于是, 当 $k \to \infty$ 时 $\|p_k\| \to 0$, 但当 $k \to \infty$ 时 $|f(p_k)| \to \infty$, 因此, f 是不连续的.

下面来考察线性算子的逆算子 (当其存在时) 的连续性.

定理 2.9-4 设 X 和 Y 为两个赋范向量空间, $A: X \to Y$ 为线性算子, 则下述两个性质等价:

(a) 线性算子 A 为单射, 算子 $A: X \to \operatorname{Im} A$ 的逆映射 $B: \operatorname{Im} A \to X$ 是连续线性算子.

(b) 存在常数 $C>0$, 使得对一切 $x \in X$,

$$\|x\| \leqslant C\|Ax\|.$$

证明 如果 (a) 成立, 则 $A: X \to \operatorname{Im} A$ 为双射 (定理 2.9-1). 因而, 对任意给定的 $x \in X$, 唯一地存在向量 $y \in \operatorname{Im} A$, 使得 $Ax=y$, 或等价地 $By=x$. 又由 B 的连续性可知存在常数 $C>0$, 使得对任何 $y \in \operatorname{Im} A, \|By\| \leqslant C\|y\|$(定理 2.9-2). 因此 (b) 成立.

如果 (b) 成立, 则由 $\operatorname{Ker} A=\{0\}$ 得 A 是单射 (定理 2.9-1), 由对所有的 $x \in X, \|x\| \leqslant C\|Ax\|$, 得到对所有的 $y \in \operatorname{Im} A$, $\|By\| \leqslant C\|y\|$. 再由定理 2.9-2, B 是连续的. □

设 X 和 Y 是同一个域 $\mathbb{K}$ 上的两个赋范线性空间. 则将域 $\mathbb{K}$ 上由所有 X 到 Y 的连续线性算子组成的空间记作

$$\mathcal{L}(X;Y) \quad \text{或当} \quad Y=X \text{ 时 } \mathcal{L}(X).$$

[15] A. A. Markoff [1889]: Sur une question posée par Mendeleieff. Izvestia Akademii Nauk SSSR 62, 1–24.

[16] 这个证明可见 E. W. Cheney [1966]: Introduction to Approximation Theory. McGraw-Hill, New York, 第 3 章第 7 节.

[17] 对 "最佳" 常数 $C(n,q)$ 的估计见:

E. Hille; C. Szegö; J. D. Tamarkin [1937]: On some generalization of a theorem of A. Markoff. Duke Mathematical Journal 8, 729–739.

在这个空间上也可以赋以范数. 关于这个范数的定义及初等性质将在下面的定理中给出.

尽管为记号简明起见不再提及, 然而下面在取上确界时向量 x 显然属于空间 X (同样的说明适用于以下多处.)

定理 2.9-5 设 X 和 Y 是两个赋范向量空间.

(a) 由

$$\|\cdot\| : A \in \mathcal{L}(X;Y) \to \|A\| := \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$$

定义的映射是向量空间 $\mathcal{L}(X;Y)$ 上的范数. 由定义, 对任何 $x \in X$,

$$\|Ax\| \leqslant \|A\| \, \|x\|.$$

(b) $A \in \mathcal{L}(X;Y)$ 的范数可以等价地定义为

$$\begin{aligned}\|A\| &= \sup_{\|x\|\leqslant 1} \|Ax\| = \sup_{\|x\|<1} \|Ax\| = \sup_{\|x\|=1} \|Ax\| \\ &= \frac{1}{r} \sup_{\|x\|\leqslant r} \|Ax\| = \frac{1}{r} \sup_{\|x\|=r} \|Ax\| \\ &= \inf\{C > 0;\ \text{对一切 } x \in X,\ \|Ax\| \leqslant C\|x\|\}.\end{aligned}$$

(c) 如果 X 是有限维空间, $A \in \mathcal{L}(X;Y)$, 则存在 $x_0 \in X$, 使得

$$x_0 \neq 0 \text{ 且 } \|A\| \, \|x_0\| = \|Ax_0\|.$$

(d) 设 Z 为赋范向量空间, 如果 $A \in \mathcal{L}(X;Y), B \in \mathcal{L}(Y;Z)$, 则 $BA \in \mathcal{L}(X;Z)$, 且

$$\|BA\| \leqslant \|A\| \, \|B\|.$$

由此, 如果 $A \in \mathcal{L}(X)$, 则对任何整数 $n \geqslant 0$,

$$\|A^n\| \leqslant \|A\|^n.$$

(e) 若 $A \in \mathcal{L}(X)$, 则 A 的任意特征值 λ 满足 $|\lambda| \leqslant \|A\|$.

证明 性质 (a), (b) 和 (d) 均可直接验证 (为证明 $\sup_{\|x\|\leqslant 1} \|Ax\| = \sup_{\|x\|<1} \|Ax\|$ 只要注意到单位球在闭单位球中稠密). 如果 X 是有限维空间, 则单位球面 $K = \{x \in X; \|x\| = 1\}$ 是紧集 (定理 2.7-1). 由于函数 $x \in X \to \|Ax\| \in \mathbb{R}$ 是连续的 (它是两个连续映射 $x \in X \to Ax \in Y, y \in Y \to \|y\| \in \mathbb{R}$ 的复合), 从而在 K 上达到其上确界. 这就证得 (c).

若 $A \in \mathcal{L}$, $Ap = \lambda p$, $p \neq 0$, 则 $\|Ap\| = |\lambda| \|p\| \leqslant \|A\| \|p\|$, 由此得证 (e). □

空间 $\mathcal{L}(X;Y)$ 上由定理 2.9-5 定义的范数称为算子范数. 在容易产生歧义时也记作

$$\|\cdot\|_{\mathcal{L}(X;Y)}.$$

在 $Y=\mathbb{K}$ 这个基本的特殊情况下, 称空间

$$X' := \mathcal{L}(X;\mathbb{K})$$

为 X 的对偶空间, 或简称为 X 的对偶. 任何 $\ell \in X'$ 的范数由

$$\|\ell\| = \sup_{x\neq 0}\frac{|\ell(x)|}{\|x\|}$$

给出.

注意, 在这个情况下, 对任意的 $\ell \in X'$ 和 $x \in X$, 也采用记号

$${}_{X'}\langle \ell, x\rangle_X := \ell(x) \text{ 或简记为 } \langle \ell, x\rangle := \ell(x).$$

在 $X=\mathbb{K}$ 的特殊情况下, 利用线性双射 $A\in\mathcal{L}(\mathbb{K};Y)\to A(1)\in Y$, 空间 $\mathcal{L}(\mathbb{K},Y)$ 可恒同于空间 Y.

有限维空间上算子范数的例子和性质将在习题 2.9-1 和 2.9-2 中给出. 当 X 为无限维空间时, 定理 2.9-5 性质 (c) 的反例由习题 2.9-4 给出. 利用连续线性泛函刻画的有限维空间的特征可见习题 2.9-5. 对偶空间的例子将会在第 3 章中出现.

习题

2.9-1 $\mathbb{K}^n$ 上的范数 $\|\cdot\|_p, 1\leqslant p\leqslant\infty$ 由定理 2.2-2 所定义. $\mathcal{L}(\mathbb{K}^n)$ 上的线性算子恒同于系数在 $\mathbb{K}$ 中的 $n\times n$ 阵 $\boldsymbol{A}=(a_{ij})$.

(1) 证明

$$\|\boldsymbol{A}\|_1 := \sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|\boldsymbol{A}\boldsymbol{x}\|_1}{\|\boldsymbol{x}\|_1} = \max_{1\leqslant j\leqslant n}\sum_{i=1}^n |a_{ij}|,$$

$$\|\boldsymbol{A}\|_2 := \sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \{\rho(\boldsymbol{A}^*\boldsymbol{A})\}^{\frac{1}{2}},$$

$$\|\boldsymbol{A}\|_\infty := \sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|\boldsymbol{A}\boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_\infty} = \max_{1\leqslant i\leqslant n}\sum_{j=1}^n |a_{ij}|,$$

这里 $\boldsymbol{A}^*$ 表示 $\boldsymbol{A}$ 的共轭阵, $\rho(\boldsymbol{B})$ 表示矩阵 $\boldsymbol{B}$ 的特征值的最大模.

注 因为向量范数 $\|\cdot\|_2$ 也被记作 $|\cdot|$, 所以在不致混淆的情况下, 上述矩阵范数 $\|\cdot\|_2$ 也可简记为 $|\cdot|$.

(2) 对下述定义的算子范数

$$\sup_{x\neq 0}\frac{\|\boldsymbol{A}\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_1},\ 1<p\leqslant\infty \text{ 和 } \sup_{x\neq 0}\frac{\|\boldsymbol{A}\boldsymbol{x}\|_\infty}{\|\boldsymbol{x}\|_p},\ 1\leqslant p<\infty,$$

寻求算子范数的公式.

2.9-2 记号同习题 2.9-1.

(1) 设 $\boldsymbol{A}$ 为实 $n\times n$ 阵, 证明: 对于 $p=1,p=2$ 和 $p=\infty$,

$$\sup\left\{\frac{\|\boldsymbol{Ax}\|_p}{\|\boldsymbol{x}\|_p};\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{x}\neq\boldsymbol{0}\right\}=\sup\left\{\frac{\|\boldsymbol{Ax}\|_p}{\|\boldsymbol{x}\|_p};\boldsymbol{x}\in\mathbb{C}^n,\boldsymbol{x}\neq\boldsymbol{0}\right\}.$$

(2) 对 $1<p<2$ 和 $2<p<\infty$, (1) 中等式是否仍然成立?

(3) 试给出一个实 $n\times n$ 阵 $\boldsymbol{A}$ 和一个 $\mathbb{C}^n$ 上的范数, 使得

$$\sup\left\{\frac{\|\boldsymbol{Ax}\|}{\|\boldsymbol{x}\|};\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{x}\neq 0\right\}<\sup\left\{\frac{\|\boldsymbol{Ax}\|}{\|\boldsymbol{x}\|};\boldsymbol{x}\in\mathbb{C}^n,\boldsymbol{x}\neq\boldsymbol{0}\right\}.$$

2.9-3 在 $\mathbb{K}^n$ 上任意给定向量范数, 在 $\mathcal{L}(\mathbb{K}^n)$ 上对任意的 $\boldsymbol{A}\in\mathcal{L}(\mathbb{K}^n)$, 由 $\|\boldsymbol{A}\|=\sup_{\boldsymbol{x}\neq\boldsymbol{0}}\frac{\|\boldsymbol{Ax}\|}{\|\boldsymbol{x}\|}$ 定义了相应的算子范数. 如同在这个问题中, 当 $\boldsymbol{A}$ 恒同于一个 $n\times n$ 阵时, 这个算子范数也被称为从属的矩阵范数 (反映它 "从属" 于给定的向量范数, 从属矩阵范数的例可见习题 2.9-1 中的 (1). 显然, 对任何从属的矩阵范数, $\rho(\boldsymbol{A})\leqslant\|\boldsymbol{A}\|$, 其中 $\rho(\boldsymbol{A})$ 表示 $\boldsymbol{A}$ 的特征值的最大模 (因为当 $\boldsymbol{Ap}=\lambda\boldsymbol{p}$ 时 $|\lambda|\,\|\boldsymbol{p}\|\leqslant\|\boldsymbol{A}\|\,\|\boldsymbol{p}\|$). 下面, 给定各系数取于 $\mathbb{K}$ 的 $n\times n$ 阵 $\boldsymbol{A}$.

(1) 证明对任意的 $\varepsilon>0$, 存在从属的矩阵范数 $\|\cdot\|$, 使得 $\|\boldsymbol{A}\|\leqslant\rho(\boldsymbol{A})+\varepsilon$.

(2) 证明当 $k\to\infty$ 时 $\boldsymbol{A}^k\to 0$, 当且仅当 $\rho(\boldsymbol{A})<1$.

(3) 证明对任意给定的从属矩阵范数, 当 $k\to\infty$ 时

$$(\|\boldsymbol{A}^k\|)^{\frac{1}{k}}\to\rho(\boldsymbol{A}).$$

(4) 证明

$$\limsup_{k\to\infty}(|\mathrm{tr}(\boldsymbol{A}^k)|^{\frac{1}{k}})=\rho(\boldsymbol{A}).$$

2.9-4 在第 2.4 节中已经给出空间 ℓ^∞ 和其上范数 $\|\cdot\|_\infty$. 用 C_0 表示所有满足 $\lim_{i\to\infty}x_i=0$ 的无穷序列 $(x_i)_{i=1}^\infty$ 的集合, 其中 $x_i\in\mathbb{K}$. 显然, c_0 是 ℓ^∞ 的一个子空间 (收敛序列必有界).

(1) 设 $\sum_{i=1}^\infty\alpha_i$ 为收敛级数, 满足对一切 $i\geqslant 1,\alpha_i>0$. 证明:

$$f:x=(x_i)_{i=1}^\infty\in(c_0;\|\cdot\|_\infty)\to f(x):=\sum_{i=1}^\infty\alpha_i x_i\in\mathbb{K}$$

是 c_0 上的连续线性泛函, $\|f\|:=\sup_{x\neq 0}\frac{|f(x)|}{\|x\|_\infty}=\sum_{i=1}^\infty\alpha_i$.

(2) 证明不存在非零向量 $x\in c_0$, 使得 $|f(x)|=\|f\|\,\|x\|_\infty$.

2.9-5 证明: 赋范向量空间 X 为有限维空间的充要条件是 X 上所有的线性泛函均是连续的.

2.9-6 (1) 设 X 和 Y 是同一个域上的两个向量空间, $A:X\to Y$ 为一个线性算子. 证明存在一个从商空间 $X/\mathrm{Ker}\ A$ 到 Y 的子空间 $\mathrm{Im}\ A$ 上的线性双射.

(2) 给定线性算子 $A:X\to Y$, 定义映射 $[A]:X/\mathrm{Ker}\ A\to Y$ 为对任何 $[x]\in X/\mathrm{Ker}\ A$, $[A][x]:=Ax$. 证明这个定义是有意义的, $[A]$ 是一个线性算子, 它是 $X/\mathrm{Ker}\ A$ 到 $\mathrm{Im}\ A$ 上的一个双射.

(3) 设 $\operatorname{Im} A$ 在 Y 中闭, 证明 $[A]$ 是连续的充要条件是 A 为连续的, 且此时 $\|[A]\|=\|A\|$.

2.9-7 (1) 设 Y 为实向量空间, $A:\mathcal{L}(\mathbb{R}^3)\to Y$ 为线性算子, 而且它在 3×3 正常正交阵集合 (此处定义为 $\mathcal{L}(\mathbb{R}^3)$ 的子集) 上的限制是常值映射, 证明 $A=0$.

(2) 对于 $n\neq 3$ 的任何维数, 上述性质是否依然成立?

2.10 赋范向量空间上的紧线性算子

本节介绍一类重要的连续线性算子.

设 X 和 Y 是同一个域上的两个赋范向量空间. 设 $A:X\to Y$ 为线性算子. 如果 X 中任何有界子集在 A 作用下的像是 Y 中的相对紧子集, 即当 B 在 X 中有界时, $\overline{A(B)}$ 是紧集, 则称 A 为紧算子.

现在我们来证明紧算子的一些初等而重要的性质. 首先, (直接) 可知任何紧线性算子均是连续的, 因而 X 到 Y 的所有紧算子的集合是空间 $\mathcal{L}(X;Y)$ 的子集 (而且是 $\mathcal{L}(X;Y)$ 的一个子空间). 第二个性质的充分性部分被经常用以证明一个线性算子的紧性. 最后两个 (也是直接的) 性质是线性算子为紧算子的充分条件.

定理 2.10-1 设 X 和 Y 是同一个域上的两个赋范向量空间, $A:X\to Y$ 为线性算子.

(a) 如果 A 是紧的, 则 A 是连续的.

(b) 算子 A 为紧算子的充要条件是给定 X 中任意的有界序列 $(x_n)_{n=1}^{\infty}$, 在 Y 中的序列 $(Ax_n)_{n=1}^{\infty}$ 有收敛子列.

(c) 如果 X 是有限维的, 则 A 是紧的.

(d) 如果 A 是连续的, 且 X 在 A 作用下的直接像 $A(X)$ 是有限维的, 则 A 是紧的.

证明 (i) 如果线性算子 $A:X\to Y$ 是紧的, X 中任何有界子集 B 的直接像 $A(B)$ 是 Y 中的有界集 (作为 Y 中紧集 $\overline{A(B)}$ 的子集; 见定理 1.13-1). 因此, A 是连续的 (定理 2.9-2(d)). (a) 得证.

(ii) 假设对任意给定的 X 中的序列 $(x_n)_{n=1}^{\infty}$, Y 中的序列 $(Ax_n)_{n=1}^{\infty}$ 有收敛子列. 再设 B 是 X 中任意的有界子集.

给定集合 $A(B)$ 中的任何序列 $(y_n)_{n=1}^{\infty}$, 取 $x_n\in B$, 使得 $y_n=Ax_n$. 于是, 序列 $(x_n)_{n=1}^{\infty}$ 是有界的, 由假设, 存在子列 $(Ax_{\sigma(n)})_{n=1}^{\infty}$ 收敛于 Y 中的元 y, 因为对所有的 $n\geqslant 1$, $y_{\sigma(n)}=Ax_{\sigma(n)}\in A(B)$, 故而

$$y=\lim_{n\to\infty} y_{\sigma(n)}\in\overline{A(B)}.$$

所以, 由定理 1.13-4, 集合 $A(B)$ 是相对紧的. 这就证得 (b) 的充分性部分.

(iii) 再证 (b) 的必要性部分. 设 A 是紧的, $(x_n)_{n=1}^{\infty}$ 是 X 中的任何有界序列.

集合 $B := \cup_{n=1}^{\infty}\{x_n\}$ 是有界的. 由假设, $\overline{A(B)}$ 是紧的. 因为对所有的 $n \geqslant 1$, $Ax_n \in A(B) \subset \overline{A(B)}$, 故而存在在 $\overline{A(B)} \subset Y$ 中收敛的子列 $(Ax_{\sigma(n)})_{n=1}^{\infty}$.

(iv) 设 X 为有限维空间, 则任何线性算子 $A : X \to Y$ 均是连续的 (定理 2.9-3), 从而它映 X 中的任何有界子集 B 为 Y 中的有界子集 $A(B)$ (定理 2.9-2(d)). 因为 $A(B) \subset A(X)$, 且 X 在 A 作用下的直接像 $A(X)$ 是 Y 的有限维子空间, 所以集合 $\overline{A(B)}$ 是紧的 (定理 2.7-1(c)). 这就证得了 (c).

(v) 设 $A \in \mathcal{L}(X;Y)$, 直接像 $A(X)$ 是有限维的, 设 B 是 X 的任何有界子集. 因为集合 $\overline{A(B)}$ 是有限维赋范线性空间 $A(X)$ 的闭的有界子集 (由 A 的连续性假设), 从而是紧的 (定理 2.7-1(c)). □

设 X 是 Y 的子空间, 记号

$$X \Subset Y$$

表示由 X 到 Y 的典则单射 (1.2 节) 为紧线性算子, 即 X 中任何有界序列包含在 Y 中收敛的子列 (定理 2.10-1(b)).

注 上述关于紧性的定义是特殊地对线性算子而言, 对非线性算子则需补充上连续性的假设 (对线性算子则自动满足, 见定理 2.10-1(a)), 见 9.12 节.

由空间 $\mathcal{C}[0,1]$ 到自身或空间 $L^2(0,1)$ 到自身的紧线性算子的简单例子将在习题 3.10-4 和 4.9-5 中给出 (这些例子之所以滞后, 是因为对它们的分析需要在后面的章节中介绍的一些概念, 如完备性). 另一个特别重要的例子是 Sobolev 空间 $H^1(\Omega)$ 到 $L^2(\Omega)$ 的典则单射.

内积空间上的紧线性算子和自伴紧算子 (4.10 节), 基于这些算子的基本谱定理 (定理 4.11-1), 在二阶线性椭圆型算子谱理论 (6.10 节) 中起着关键性的作用.

习题

2.10-1 设 X 和 Y 是同一个域上的两个赋范向量空间, $A : X \to Y$ 为紧线性算子, 证明: 空间 $\operatorname{Im} A$ 是可分的.

2.10-2 设 X 为 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的赋范向量空间, $A : X \to X$ 为紧线性算子, 又设 $\lambda \in \mathbb{K}, \lambda \neq 0$. 证明 $(\lambda I - A)$ 是单射的充要条件是当 $\|x\| \to \infty$ 时 $\|(\lambda I - A)x\| \to \infty$[18].

---

[18] 这个性质的发现见:

G. Dinca [2001]: A Fredholm-type result for a couple of nonlinear operators, Comptes Rendus de l'Académie des Sciences de Paris, Série 1, 333, 4015–4019.

2.11 赋范向量空间上的连续多重线性映射; 空间 $\mathcal{L}_k(X_1, X_2, \cdots, X_k; Y)$

在本节中, k 为一个 $\geqslant 2$ 的整数, $X_\ell, 1 \leqslant \ell \leqslant k$ 和 Y 表示同一个域 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的向量空间.

乘积空间

$$X_1 \times X_2 \times \cdots \times X_k$$

为所有形如 $(x_1, x_2, \cdots, x_k)$ 的元素的集合, 其中 $x_\ell \in X_\ell, 1 \leqslant \ell \leqslant k$. 在乘积空间上定义加法和数乘为

$$(x_1, x_2, \cdots, x_k) + (y_1, y_2, \cdots, y_k) := (x_1 + y_1, x_2 + y_2, \cdots, x_k + y_k),$$
$$\alpha(x_1, x_2, \cdots, x_k) := (\alpha x_1, \alpha x_2, \cdots, \alpha x_k),$$

则 $X_1 \times X_2 \times \cdots \times X_k$ 为 $\mathbb{K}$ 上的向量空间.

称映射 $A : X_1 \times X_2 \times \cdots \times X_k \to Y$ 为由 $X_1 \times X_2 \times \cdots \times X_k$ 到 Y 的*多重线性*或 *k 线性映射*, 是指对每个变量 $x_\ell \in X_\ell, 1 \leqslant \ell \leqslant k$, 当其他 $k-1$ 个变量固定时是线性的. 当 $k=2$ 或 $k=3$ 时多重线性映射称为*双线性的*或*三线性的*. 如果 $Y = \mathbb{K}$, 则多重线性映射也被称为*多重线性泛函*或*多线性型*.

当然, 必须把这个概念与乘积空间 $X_1 \times X_2 \times \cdots \times X_k$ 到 Y 的线性映射的概念区分开来. 例如, 当 $k=2$ 时, 由 $X_1 \times X_2$ 到 Y 的线性映射满足 (记号自明)

$$A((x_1, x_2) + (y_1, y_2)) = A(x_1, x_2) + A(y_1, y_2),$$
$$A(\alpha(x_1, x_2)) = \alpha A(x_1, x_2),$$

而 $X_1 \times X_2$ 到 Y 的双线性映射则满足

$$A((x_1, x_2) + (y_1, y_2)) = A(x_1, x_2) + A(y_1, y_2) + A(x_1, y_2) + A(x_2, y_1),$$
$$A(\alpha(x_1, x_2)) = \alpha^2 A(x_1, x_2).$$

定义映射的加法和数乘为 (记号自明)

$$(A + B)(x_1, x_2, \cdots, x_k) := A(x_1, x_2, \cdots, x_k) + B(x_1, x_2, \cdots, x_k),$$
$$(\alpha A)(x_1, x_2, \cdots, x_k) := \alpha(A(x_1, x_2, \cdots, x_k)),$$

则由所有 $X_1 \times X_2 \times \cdots \times X_k$ 到 Y 的多重线性映射组成的集合成为 $\mathbb{K}$ 上的一个向量空间.

用 $\mathfrak{S}_k$ 表示由集合 $\{1, 2, \cdots, k\}$ 的置换全体组成的集合. 如果 $X_1 = X_2 = \cdots = X_k = X, A$ 为 $X \times X \times \cdots \times X$ (k 个因子) 到 Y 的多重线性映射. 如果对一

切 $x_\ell \in X_\ell, 1 \leqslant \ell \leqslant k$ 和所有的 $\sigma \in \mathfrak{S}_k$,

$$A(x_{\sigma(1)}, x_{\sigma(2)}, \cdots, x_{\sigma(k)}) = A(x_1, x_2, \cdots, x_k),$$

则称 A 为对称的, 如果

$$A(x_{\sigma(1)}, x_{\sigma(2)}, \cdots, x_{\sigma(k)}) = \varepsilon(\sigma) A(x_1, x_2, \cdots, x_k),$$

其中 $\varepsilon(\sigma)$ 表示 σ 的符号, 则称 A 为交错的. 系数在 $\mathbb{K}$ 上的 $k \times k$ 矩阵的行列式, 作为矩阵列向量的函数, 就是一个 $X = \mathbb{K}^k$ 的交错多重泛函的例子.

当空间 $X_1, X_2, \cdots, X_k$ 和 Y 均为赋范向量空间, 在乘积空间 $X_1 \times X_2 \times \cdots \times X_k$ 上赋以乘积拓扑 (2.2 节) 时, 连续多重线性映射具有特殊的性质. 下面的定理罗列了这类算子的各种特征, 它们可以视为在定理 2.9-2 中建立的连续线性算子的特征的"多重线性的推广".

定理 2.11-1 设 $X_\ell, 1 \leqslant \ell \leqslant k$ 和 Y 均为赋范向量空间,

$$A : X := X_1 \times X_2 \times \cdots \times X_k \to Y$$

为多重线性映射, 则下列性质等价:

(a) 映射 $A : X \to Y$ 是连续的.

(b) 映射 A 在原点处连续.

(c) 存在常数 $C > 0$, 使得对任意的 $x = (x_1, x_2, \cdots, x_k) \in X$,

$$\|Ax\|_Y \leqslant C \|x_1\|_{X_1} \|x_2\|_{X_2} \cdots \|x_k\|_{X_k}.$$

(d) X 中任何有界子集的直接像是 Y 的有界子集.

证明 在证明中, 不失一般性, 我们总假设 $X_1 \times X_2 \times \cdots \times X_k$ 上的乘积拓扑由范数 $|||\cdot|||_\infty$ 导出, 其定义为对 $x = (x_1, x_2, \cdots, x_k) \in X$,

$$|||x|||_\infty = \max_{1 \leqslant \ell \leqslant k} \|x_\ell\|_{X_\ell}.$$

显然, 由 (a) 即可导出 (b). 如果 (b) 成立, 则 Y 中的闭单位球在 A 下的逆像包含一个以 X 的原点为中心的闭球, 记其半径为 $\alpha > 0$. 由范数 $|||\cdot|||_\infty$ 的定义可知, 当向量 $\widetilde{x} = (\widetilde{x_1}, \widetilde{x_2}, \cdots, \widetilde{x_k}) \in X$ 满足 $\|\widetilde{x_\ell}\|_{X_\ell} \leqslant \alpha$ 时, $\|A\widetilde{x}\|_Y \leqslant 1$.

给定任意向量 $x = (x_1, x_2, \cdots, x_k) \in X$, 满足 $x_\ell \neq 0, 1 \leqslant \ell \leqslant k$ (否则, (c) 中不等式自然成立), 令 $\widetilde{x} = (\widetilde{x_1}, \widetilde{x_2}, \cdots, \widetilde{x_\ell})$, 其中 $\widetilde{x_\ell} := \alpha(\|x_\ell\|_{X_\ell})^{-1} x_\ell, 1 \leqslant \ell \leqslant k$. 则对 $1 \leqslant \ell \leqslant k$, $\|\widetilde{x_\ell}\|_{X_\ell} = \alpha$, 从而 $\|A\widetilde{x}\|_Y \leqslant 1$. 由于假设 A 是多重线性的, 故

$$Ax = \frac{1}{\alpha^k} \|x_1\|_{X_1} \|x_2\|_{X_2} \cdots \|x_\ell\|_{X_\ell} A\widetilde{x},$$

取 $C := \frac{1}{\alpha^k}$, (c) 中的不等式成立.

再设 (c) 成立. 因为 X 中的任何有界子集 B 均可包含于一个以 X 的原点为中心、半径 $r = r(B) > 0$ 的球中, 所以直接像 $A(B)$ 包含于以 Y 的原点为中心、半径为 Cr^k 的球中, 这就证得 $A(B)$ 是有界的. 于是由 (c) 得到 (d).

如果 (d) 成立, X 中闭单位球的直接像在 Y 中有界. 由范数 $|||\cdot|||_\infty$ 的定义, 存在常数 $C > 0$, 使得 $\|x_\ell\|_{X_\ell} \leqslant 1$, 当 $1 \leqslant \ell \leqslant k$ 时, $\|A(x_1, x_2, \cdots, x_k)\|_Y \leqslant C$. 于是, 由 A 的多重线性, 对一切 $x = (x_1, x_2, \cdots, x_k) \in X$,

$$\|Ax\|_Y \leqslant C\|x_1\|_{X_1}\|x_2\|_{X_2}\cdots\|x_k\|_{X_k}.$$

给定 $x = (x_1, x_2, \cdots, x_k) \in X$ 和 $a = (a_1, a_2, \cdots, a_k) \in X$, 由 A 的多重线性, 差 $A(x) - A(a)$ 可以表示为

$$\begin{aligned} A(x) - A(a) = &A(x_1 - a_1, x_2, x_3, \cdots, x_k) + \\ &A(a_1, x_2 - a_2, x_3, \cdots, x_k) + \\ &\vdots \\ &A(a_1, a_2, \cdots, a_{k-1}, x_k - a_k). \end{aligned}$$

因此,

$$\begin{aligned} \|A(x) - A(a)\|_Y \leqslant C(&\|x_1 - a_1\|_{X_1}\|x_2\|_{X_2}\cdots\|x_k\|_{X_k} + \\ &\|a_1\|_{X_1}\|x_2 - a_2\|_{X_2}\cdots\|x_k\|_{X_k} + \\ &\vdots \\ &\|a_1\|_{X_1}\|a_2\|_{X_2}\cdots\|x_k - a_k\|_{X_k}). \end{aligned}$$

令 $M := |||a|||_\infty$, $\delta := |||x - a|||_\infty$. 因为 $|||x|||_\infty \leqslant M + \delta$, 由前述不等式即得

$$\|A(x) - A(a)\|_Y \leqslant C\delta\{(M+\delta)^{k-1} + M(M+\delta)^{k-2} + \cdots + M^{k-1}\}.$$

固定 a, 当 $\delta = |||x - a|||_\infty$ 趋于 0 时这个不等式的右边趋于 0, 从而映射 A 是连续的. 这就由 (d) 导出了 (a). □

比较一下下面两个不等式是有意义的. 对所有的 $x = (x_1, \cdots, x_k) \in X$,

$$\|Ax\|_Y \leqslant C\|x_1\|_{X_1}\|x_2\|_{X_2}\cdots\|x_k\|_{X_k},$$

即多重线性映射 $A: X_1 \times X_2 \times \cdots \times X_k \to Y$ 为连续的. 对所有的 $x = (x_1, \cdots, x_k) \in X$,

$$\|Ax\|_Y \leqslant C(\|x_1\|_{X_1} + \|x_2\|_{X_2} + \cdots + \|x_k\|_{X_k})$$

(或 $\|Ax\|_Y \leqslant C\max_{1\leqslant\ell\leqslant k}\|x_\ell\|_{X_\ell}$) 意为线性映射 $A: X_1 \times X_2 \times \cdots \times X_k \to Y$ 为连续线的.

给定向量空间 X, 因为 $\|\lambda x\| = |\lambda|\,\|x\|$, 故而映射 $(\alpha, x) \in \mathbb{K} \times X \to \alpha x \in X$ 就是一个连续双线性映射的例子.

给定 $1 \leqslant p \leqslant \infty$, 设 q 为 p 的共轭指数. 双线性映射

$$((x_i)_{i=1}^{\infty}, (y_i)_{i=1}^{\infty}) \in \ell^p \times \ell^q \to (x_i y_i)_{i=1}^{\infty} \in \ell^{\infty}$$

和

$$((x_i)_{i=1}^{\infty}, (y_i)_{i=1}^{\infty}) \in \ell^p \times \ell^q \to \sum_{i=1}^{\infty} x_i y_i \in \mathbb{K},$$

由 Hölder 不等式 (定理 2.4-1) 也提供了连续双线性映射的例子.

注　对任何 $1 \leqslant p \leqslant \infty$, 卷积 $(f, g) \in L^1(\mathbb{R}^n) \times L^p(\mathbb{R}^n) \to f * g \in L^p(\mathbb{R}^n)$ 提供了连续双线性映射的另一个例子 (习题 2.6-4).

下面的结果推广了线性映射的一个性质 (定理 2.9-3(b)), 同时也提供了连续的多重线性映射的其他例子.

定理 2.11-2　设 $X_1, X_2, \cdots, X_k$ 均为有限维空间, Y 为赋范向量空间, 则多重线性映射 $A : X_1 \times X_2 \times \cdots \times X_k \to Y$ 必是连续的.

证明　对每个 $1 \leqslant \ell \leqslant k$, 设 $(e_{i(\ell)}^{\ell})_{i(\ell)=1}^{m(\ell)}$ 为 X_ℓ 的基. 给定向量 $x = (x_1, x_2, \cdots, x_k) \in X_1 \times X_2 \times \cdots \times X_k$, 其中 $x_\ell = \sum_{i(\ell)=1}^{m(\ell)} x_{i(\ell)}^{\ell} e_{i(\ell)}^{\ell}, 1 \leqslant k \leqslant \ell$, 由 A 的多重线性即得

$$A(x) = \sum_{i(1)=1}^{m(1)} \sum_{i(2)=1}^{m(2)} \cdots \sum_{i(k)=1}^{m(k)} x_{i(1)}^1 x_{i(2)}^2 \cdots x_{i(k)}^k A(e_{i(1)}^1, e_{i(2)}^2, \cdots, e_{i(k)}^k).$$

因为这个关系式中右边的和是有限的, 因而存在常数 C, 使得 (这里为确定起见, 对每个空间 $X_j, 1 \leqslant j \leqslant k$ 均赋以范数 $\|\cdot\|_\infty$)

$$\|A(x)\|_Y \leqslant C\|x_1\|_\infty \|x_2\|_\infty \cdots \|x_k\|_\infty.$$

因为有限维空间上所有的范数均等价 (定理 2.7-1(a)), 由定理 2.11-1(c) 即得所证之结论. □

例如, 定理 2.11-2 说明系数在 $\mathbb{K}$ 中的 $k \times k$ 阵的行列式是 $\mathbb{K}^k$ 到 $\mathbb{K}$ 的连续多重线性泛函.

相反地, 连续线性算子必为一致连续的性质 (定理 2.9-3(a)) 则不能扩充到多重线性映射.

定理 2.11-3　非零的连续多重线性映射不是一致连续的.

证明　给定非零的连续多重线性映射 $A : X_1 \times X_2 \times \cdots \times X_k \to Y$, 设 $a \in$

$X, A(a) \neq 0$. 对每个整数 $n \geqslant 1$, 令

$$x_n := na, \quad y_n := \left(n - \frac{1}{n^{k-1}}\right) a,$$

于是当 $n \to \infty$ 时 (由假设 $k \geqslant 2$)

$$\|x_n - y_n\|_X = \frac{1}{n^{k-1}} \|a\|_X \to 0.$$

由 A 的多重线性

$$\|A(x_n) - A(y_n)\|_Y = \left(n^k - \left(n - \frac{1}{n^{k-1}}\right)^k\right) \|A(a)\|_Y,$$

又由 $\lim_{n\to\infty}(n^k - (n - \frac{1}{n^{k-1}})^k) = k$, 故而存在 $n_0 \geqslant 1$, 使得当 $n \geqslant n_0$ 时

$$\|A(x_n) - A(y_n)\|_Y \geqslant \frac{k}{2} \|A(a)\|_Y,$$

所以 A 不是一致连续的. □

设 $X_1, X_2, \cdots, X_k$ 和 Y 是同一个域 $\mathbb{K}$ 上的赋范向量空间. 由所有从 $X_1 \times X_2 \times \cdots \times X_k$ 到 Y 的连续多重线性映射组成的域 $\mathbb{K}$ 上的向量空间记作

$$\mathcal{L}_k(X_1, X_2, \cdots, X_k; Y) \text{ 或 } \mathcal{L}_k(X_1 \times X_2 \times \cdots \times X_k; Y)$$
$$\text{或 } X_\ell = X, \text{ 当 } 1 \leqslant \ell \leqslant k \text{ 时, } \mathcal{L}_k(X; Y).$$

这个空间也能赋以范数. 这个范数的定义及初等性质将在下一个定理中给出, 它们可视为定理 2.9-5 的 (a) 和 (b) 的"多重线性的推广".

定理 2.11-4 设 $X_1, X_2, \cdots, X_k$ 和 Y 为赋范向量空间.

(a) 由

$$\|\cdot\| : A \in \mathcal{L}_k(X_1, X_2, \cdots, X_k; Y) \to \|A\| := \sup_{\substack{x_\ell \neq 0 \\ 1 \leqslant \ell \leqslant k}} \frac{\|A(x_1, x_2, \cdots, x_k)\|_Y}{\|x_1\|_{X_1} \|x_2\|_{X_2} \cdots \|x_k\|_{X_k}}$$

定义的映射是向量空间 $\mathcal{L}_k(X_1, X_2, \cdots, X_k; Y)$ 上的范数. 由定义, 对任何 $(x_1, x_2, \cdots, x_k) \in X_1 \times X_2 \times \cdots \times X_k$,

$$\|A(x_1, x_2, \cdots, x_k)\|_Y \leqslant \|A\| \, \|x_1\|_{X_1} \|x_2\|_{X_2} \cdots \|x_k\|_{X_k}.$$

(b) 任何 $A \in \mathcal{L}_k(X_1, X_2, \cdots, X_k; Y)$ 的范数可以等价地定义为

$$\begin{aligned} \|A\| &= \sup_{\substack{\|x_\ell\|_{X_\ell} \leqslant 1 \\ 1 \leqslant \ell \leqslant k}} \|A(x_1, x_2, \cdots, x_k)\|_Y \\ &= \sup_{\substack{\|x_\ell\|_{X_\ell} = 1 \\ 1 \leqslant \ell \leqslant k}} \|A(x_1, x_2, \cdots, x_k)\|_Y \end{aligned}$$

或者

$$\|A\| = \inf\{C \geqslant 0;\ \text{对任何}\ (x_1, x_2, \cdots, x_k) \in X_1 \times X_2 \times \cdots \times X_k, \|A(x_1, x_2, \cdots, x_k)\|_Y \\ \leqslant C\|x_1\|_{X_1}\|x_2\|_{X_2}\cdots\|x_k\|_{X_k}\}.$$

证明 性质 (a) 和 (b) 均可由定理 2.11-1(c) 和多重线性映射的定义直接导出. □

我们将在第 7 章中, 由阶数 $\geqslant 2$ 的 Fréchet 导数给出连续的多重线性映射的基本例子. 下面的例子对这类导数的研究特别有用, 它说明任何的连续多重线性映射的空间可以由连续线性映射的空间迭代得到.

定理 2.11-5 设 $X_1, X_2, \cdots, X_k$ 和 Y 均为赋范向量空间. 则存在线性的等距双射

$$\iota : \mathcal{L}(X_1; \mathcal{L}(X_2; \cdots ; \mathcal{L}(X_k; Y)) \cdots) \to \mathcal{L}_k(X_1, X_2, \cdots, X_k; Y).$$

证明 我们对 $k = 2$ 的情况做出证明, 即证明对给定的赋范向量空间 X_1, X_2 和 Y, 存在线性的等距双射

$$\iota : \mathcal{L}(X_1; \mathcal{L}(X_2; Y)) \to \mathcal{L}_2(X_1, X_2; Y).$$

为简明起见, 所有的范数都用同一个记号 $\|\cdot\|$.

给定任意的 $A \in \mathcal{L}(X_1; \mathcal{L}(X_2; Y))$, 定义映射 $\widetilde{A} : X_1 \times X_2 \to Y$ 为对所有的 $(x_1, x_2) \in X_1 \times X_2$,

$$\widetilde{A}(x_1, x_2) = (Ax_1)x_2,$$

则 $\widetilde{A}$ 显然是双线性的. $\widetilde{A}$ 又是连续的, 这是因为对任意的 $(x_1, x_2) \in X_1 \times X_2$,

$$\|\widetilde{A}(x_1, x_2)\| = \|(Ax_1)x_2\| \leqslant \|Ax_1\|\,\|x_2\| \leqslant \|A\|\,\|x_1\|\,\|x_2\|,$$

由此 $\|\widetilde{A}\| \leqslant \|A\|$. 为证 $\|\widetilde{A}\| \geqslant \|A\|$, 任意给定 $\varepsilon > 0$. 由 $\|A\|$ 的定义, 存在 $\widetilde{x_1} \in X_1$, 使 $\|\widetilde{x_1}\| = 1$ 且 $\|A\widetilde{x_1}\| \geqslant \|A\|(1-\varepsilon)$. 对这个确定的 $\widetilde{x_1} \in X_1$, 同样地由 $\|A\widetilde{x_1}\|$ 的定义, 存在 $\widetilde{x_2} \in X_2$, 使 $\|\widetilde{x_2}\| = 1$, 且 $\|(A\widetilde{x_1})\widetilde{x_2}\| \geqslant \|A\widetilde{x_1}\|(1-\varepsilon)$.

因此,

$$\begin{aligned}\|\widetilde{A}\| = \sup_{\|x_1\|=\|x_2\|=1} \|\widetilde{A}(x_1, x_2)\| &\geqslant \|\widetilde{A}(\widetilde{x_1}, \widetilde{x_2})\| \\ &= \|(A\widetilde{x_1})\widetilde{x_2}\| \geqslant \|A\|(1-\varepsilon)^2.\end{aligned}$$

由 $\varepsilon > 0$ 的任意性, 即得 $\|\widetilde{A}\| \geqslant \|A\|$. 这就证得

$$\|\widetilde{A}\| = \|A\|.$$

还要证明由

$$\iota : A \in \mathcal{L}(X_1; \mathcal{L}(X_2; Y)) \to \iota(A) := \widetilde{A} \in \mathcal{L}_2(X_1, X_2; Y)$$

定义的线性等距算子是一个满射. 给定任何 $\widetilde{B} \in \mathcal{L}_2(X_1, X_2; Y)$, 对每个 $x_1 \in X_1$, 定义映射 $Bx_1 : X_2 \to Y$ 为

$$(Bx_1)x_2 := \widetilde{B}(x_1, x_2), \quad x_2 \in X_2.$$

于是, 对每个 $x_1 \in X_1$, 映射 $Bx_1 : X_2 \to Y$ 显然是线性的, 又因为对任何 $x_2 \in X_2$,

$$\|(Bx_1)x_2\| = \|\widetilde{B}(x_1, x_2)\| \leqslant (\|\widetilde{B}\| \, \|x_1\|)\|x_2\|,$$

所以也是连续的, 而且 $\|Bx_1\| \leqslant \|\widetilde{B}\| \, \|x_1\|$. 这样, 对每个 $x_1 \in X_1$, 有 $Bx_1 \in \mathcal{L}(X_2; Y)$. 这样定义的映射 $B : x_1 \in X_1 \to Bx_1 \in \mathcal{L}(X_2; Y)$ 显然是线性的, 由 $\|Bx_1\| \leqslant \|\widetilde{B}\| \, \|x_1\|$, 它也是连续的, 因此, $B \in \mathcal{L}(X_1; \mathcal{L}(X_2; Y))$. 由 $\iota(B) = \widetilde{B}$ 即知 ι 是满射.

$k \geqslant 3$ 的情况可类似证明, 故而略去. □

例如, 给定两个赋范向量空间 X 和 Y, 由定理 2.11-5, 可以把空间 $\mathcal{L}(X; \mathcal{L}(X; Y))$ 与 $\mathcal{L}_2(X; Y)$ 视为同一, 更一般地, 对任何整数 $k \geqslant 2$, 可以把空间 $\mathcal{L}(X; \mathcal{L}_{k-1}(X; Y))$ 和 $\mathcal{L}_k(X; Y)$ 视为同一.

习题

2.11-1 设 Ω 为 $\mathbb{R}^n$ 的开子集, $1 < p_j < \infty$, $1 \leqslant j \leqslant m$ 满足 $\frac{1}{s} := \sum_{j=1}^m \frac{1}{p_j} \leqslant 1$. 证明多重线性映射

$$(f_1, f_2, \cdots, f_m) \in L^{p_1}(\Omega) \times L^{p_2}(\Omega) \times \cdots \times L^{p_m}(\Omega) \to \prod_{j=1}^m f_j \in L^s(\Omega)$$

有定义且是连续的.

2.11-2 设 X 和 Y 是同一个域上的两个赋范向量空间, A 是由 $X \times X \times \cdots \times X$ (k 个因子) 到 Y 的多重线性对称的连续映射, 证明存在常数 $C(k)$[19], 使得

$$\|A\| \leqslant C(k) \sup_{\|x\| \leqslant 1} \|A(x, x, \cdots, x)\|.$$

2.12 Korovkin 定理

我们将在后面一节中导出 $\mathbb{R}$ 中紧区间上的连续函数可以用多项式或三角多项式逼近. 下面的 “抽象” 的定理将应用于给出上述重要结果的简单证明.

回顾一下, 给定紧的距离空间 (K, d), 记所有连续函数 $f : K \to \mathbb{R}$ 组成的空间为 $C(K)$, 其上赋以 sup 范数, 即 $\|f\| = \sup_{x \in K} |f(x)|$ (2.3 节).

[19] 可以证明最佳常数 $C(k)$ 为 $\frac{k^k}{k!}$, 例如可见:

L. Nachbin [1969]: Topology on Spaces of Holomorphic Mappings. Springer, Berlin.

定理 2.12-1 (Korovkin 定理[20]) 设 (K,d) 为紧距离空间, 函数 $\varphi \in \mathcal{C}[0,\infty[$ 满足对所有的 $t>0, \varphi(t)>0$, 对每个 $x \in K$, 函数 $\psi_x \in \mathcal{C}(K)$ 定义为

$$\psi_x(y) := \varphi(d(x,y)), \quad y \in K.$$

设线性算子 $A_n : \mathcal{C}(K) \to \mathcal{C}(K)$ 的序列 $(A_n)_{n=0}^{\infty}$ 满足下述两个性质. 第一, 对每个 $n \geqslant 0, A_n$ 为保持非负的, 即如果 $f \in \mathcal{C}(K)$, 且对一切 $x \in K, f(x) \geqslant 0$, 则对一切 $x \in K$,

$$A_n f(x) \geqslant 0.$$

第二, 对定义为对一切 $x \in K, f_0(x) = 1$ 的函数 $f_0 \in \mathcal{C}(K)$,

$$\lim_{n\to\infty} \|f_0 - A_n f_0\| = 0.$$

第三,

$$\lim_{n\to\infty} \left(\sup_{x\in K} |(A_n \psi_x)(x)| \right) = 0,$$

则对一切 $f \in \mathcal{C}(K)$, 有

$$\lim_{n\to\infty} \|f - A_n f\| = 0.$$

证明 关于证明中涉及的连续性与紧的有关性质, 读者可参阅第 1 章. 我们的目标是要证明对任意给定的函数 $f \in \mathcal{C}(K)$ 和 $\varepsilon > 0$, 存在 $n_0 = n_0(f,\varepsilon) \geqslant 0$, 使得当 $n \geqslant n_0$ 时

$$\sup_{x\in K} |(A_n f)(x) - f(x)| \leqslant \varepsilon.$$

以下, 给定 $f \in \mathcal{C}(K)$ 和 $\varepsilon > 0$.

(i) 一个技术性的不等式: 存在常数 $C = C(f,\varepsilon)$, 使得对任意的 $x, y \in K$, 均有

$$|f(y) - f(x)| \leqslant \tilde{\varepsilon} + 2C\|f\|\psi_x(y),$$

其中 (注意到因为序列 $(A_n f_0)_{n=0}^{\infty}$ 在 $\mathcal{C}(K)$ 中收敛, 故而 $\sup_{n\geqslant 0} \|A_n f_0\| < \infty$)

$$\tilde{\varepsilon} := \frac{\varepsilon}{2\sup\limits_{n\geqslant 0} \|A_n f_0\|} > 0.$$

因为函数 $f \in \mathcal{C}(K)$ 是一致连续的 (集合 K 是紧的), 所以存在 $\delta = \delta(f,\varepsilon) > 0$, 使得 $x, y \in K$ 满足当 $d(x,y) < \delta$ 时

$$|f(y) - f(x)| \leqslant \tilde{\varepsilon}.$$

[20] P. P. Korovkin [1959]: On convergence of linear positive operators in the space of continuous functions. Doklady Akademii Nauk SSSR 90, 961–964 (俄文).

因此, 只要讨论 $x, y \in K$ 满足 $d(x,y) \geqslant \delta$ 的情况. 函数 $d : K \times K \to \mathbb{R}$ 是连续的 (因为 $|d(\tilde{x},\tilde{y}) - d(x,y)| \leqslant d(\tilde{x},x) + d(\tilde{y},y)$, 这个不等式右边即定义了 $K \times K$ 上诱导出乘积拓扑的一个距离), 乘积空间 $K \times K$ 是紧的. 作为闭区间 $[\delta, \infty[$ 在连续函数 d 下的逆像

$$\mathcal{K} := \{(x,y) \in K \times K; d(x,y) \geqslant \delta\}$$

是闭的, 从而在 $K \times K$ 中是紧的.

复合函数 $\varphi \circ d : \mathcal{K} \to \mathbb{R}$ 是连续的, 且在 $\mathcal{K}$ 上 > 0 (由假设, φ 是连续的, 且在 $]0, \infty[$ 上 > 0), 因而在 $\mathcal{K}$ 上达到其最小值. 令 $C = C(\varphi, \delta) = C(\varphi, f, \varepsilon) > 0$ 定义为

$$C := \frac{1}{\inf\limits_{(x,y)\in\mathcal{K}} \varphi(d(x,y))}.$$

因为当 $x, y \in K$ 时 $\psi_x(y) = \varphi(d(x,y))$, 所以, 当 $x, y \in K$, 且当 $d(x,y) \geqslant \delta$ 时,

$$C\psi_x(y) \geqslant 1.$$

这样, 当 $x, y \in K$, 且当 $d(x,y) \geqslant \delta$ 时

$$|f(y) - f(x)| \leqslant 2\|f\| \leqslant 2C\|f\|\psi_x(y).$$

综合以上对 $d(x,y) \leqslant \delta$ 和 $d(x,y) \geqslant \delta$ 两种情况的讨论, 即得所要证明的技术上的不等式.

(ii) 对每个 $x \in K$, (i) 中得到的技术上的不等式可重新表达为

$$-\tilde{\varepsilon} f_0 - 2C\|f\|\psi_x \leqslant f - f(x) f_0 \leqslant \tilde{\varepsilon} f_0 + 2C\|f\|\psi_x.$$

由线性算子 A_n 保持非负性的假设, 对任何 $n \geqslant 0$, 得

$$-\tilde{\varepsilon} A_n f_0 - 2C\|f\|A_n\psi_x \leqslant A_n f - f(x) A_n f_0 \leqslant \tilde{\varepsilon} A_n f_0 + 2C\|f\|A_n\psi_x,$$

或等价地, 对任何 $y \in K$,

$$|(A_n f)(y) - f(x)[(A_n f_0)(y)]| \leqslant \tilde{\varepsilon}(A_n f_0)(y) + 2C\|f\|A_n\psi_x(y).$$

在这个不等式中令 $y = x$, 即得对任何 $x \in K$ 和 $n \geqslant 0$,

$$|(A_n f)(x) - f(x)[(A_n f_0)(x)]| \leqslant \tilde{\varepsilon}(A_n f_0)(x) + 2C\|f\|A_n\psi_x(x).$$

(iii) 由后一个不等式可知, 对任何 $x \in K$ 和 $n \geqslant 0$,

$$\begin{aligned}|(A_n f)(x) - f(x)| &\leqslant |(A_n f)(x) - f(x)[(A_n f_0)(x)]| + |f(x)[(A_n f_0 - f_0)(x)]| \\ &\leqslant \tilde{\varepsilon}(A_n f_0)(x) + 2C\|f\|A_n\psi_x(x) + \|f\|\, |(A_n f_0 - f_0)(x)|.\end{aligned}$$

由 (i) 中 $\tilde{\varepsilon}$ 的定义可知, 对任何 $x \in K$ 和 $n \geqslant 0$,

$$\tilde{\varepsilon}(A_n f_0)(x) \leqslant \tilde{\varepsilon}\|A_n f_0\| \leqslant \frac{\varepsilon}{2}.$$

另一方面, 对任何 $x \in K$ 和 $n \geqslant 0$,

$$\begin{aligned}&2C\|f\|(A_n\psi_x)(x) + \|f\|\,|(A_n f_0 - f_0)(x)|\\ &\leqslant 2C\|f\|\sup_{x\in K}|(A_n\psi_x)(x)| + \|f\|\,\|A_n f_0 - f_0\|.\end{aligned}$$

这样, 由对算子 A_n 的假设可知, 存在 $n_0 = n_0(f, \varepsilon)$, 使得当 $n \geqslant n_0$ 时

$$2C\|f\|\sup_{x\in K}|(A_n\psi_x)(x)| + \|f\|\,\|A_n f_0 - f_0\| \leqslant \frac{\varepsilon}{2},$$

此即是要证明的. □

注 (1) 并不假设线性算子 $A_n : \mathcal{C}(K) \to \mathcal{C}(K)$ 是连续的.

(2) 函数 $\varphi \in \mathcal{C}[0, \infty[$ 必定满足 $\varphi(0) = 0$. 这是因为由定理 2.12-1, 对每个 $x \in K$ 有 $\lim_{n\to\infty}\|A_n\psi_x - \psi_x\| = 0$. 特别地, 有 $\lim_{n\to\infty}(A_n\psi_x)(x) = \psi_x(x) = \varphi(d(x,x)) = \varphi(0)$, 而由假设, $\lim_{n\to\infty}|(A_n\psi_x)(x)| = 0$.

(3) 保持非负性的线性算子称为单调算子, 这是一个潜在容易误解的术语, 因为逆阵的元素均非负的矩阵称为单调矩阵. 事实上, “单调算子” 常指一类特殊的非线性算子, 在 9.13 节中将要介绍.

2.13 Korovkin 定理对多项式逼近的应用; Bohman 定理、Bernstein 定理和 Weierstrass 定理

Korovkin 定理的第一个应用是对赋以 sup 范数的空间 $\mathcal{C}[0,1]$.

定理 2.13-1 (Bohman 定理[21]) 设有线性算子 $A_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ 的序列 $(A_n)_{n=0}^{\infty}$, 满足以下的两个性质. 第一, 对每个 $n \geqslant 0$, A_n 均保持非负性, 即对 $f \in \mathcal{C}[0,1]$, 若当 $0 \leqslant x \leqslant 1$ 时 $f(x) \geqslant 0$, 则当 $0 \leqslant x \leqslant 1$ 时

$$(A_n f)(x) \geqslant 0.$$

第二, 对于 $f_p \in \mathcal{C}[0,1], p = 0,1,2$, 其中

$$f_0(x) = 1, \quad f_1(x) = x, \quad f_2(x) = x^2, \quad 0 \leqslant x \leqslant 1,$$
$$\lim_{n\to\infty}\|f_p - A_n f_p\| = 0 \quad p = 0,1,2.$$

成立,

[21] H. Bohman [1952]: On approximation of continous and of analytic functions. Arkiv för Mathematik 2, 43–56.

则对任何 $f \in \mathcal{C}[0,1]$, 均有

$$\lim_{n\to\infty} \|f - A_n f\| = 0.$$

证明 我们将证明 Korovkin 定理 (定理 2.12-1) 可应用于函数 $\varphi \in \mathcal{C}[0,\infty[$, 其中当 $t \geqslant 0$ 时 $\varphi(t) = t^2$. 为此, 唯一需要验证的是

$$\lim_{n\to\infty} \left(\sup_{0\leqslant x\leqslant 1} |(A_n\psi_x)(x)| \right) = 0,$$

这里当 $0 \leqslant x \leqslant 1$ 时 $\psi_x \in \mathcal{C}[0,1]$ 定义为当 $0 \leqslant y \leqslant 1$ 时

$$\begin{aligned} \psi_x(y) := \varphi(|x-y|) &= |x-y|^2 \\ &= x^2 f_0(y) - 2x f_1(y) + f_2(y) \end{aligned}$$

或等价地表示为 $\psi_x := x^2 f_0 - 2x f_1 + f_2$. 结合 $0 \leqslant x \leqslant 1$ 的关系式

$$\begin{aligned} &A_n\psi_x = x^2 A_n f_0 - 2x A_n f_1 + A_n f_2, \\ &x^2 f_0(x) - 2x f_1(x) + f_2(x) = 0, \end{aligned}$$

即得当 $0 \leqslant x \leqslant 1$ 时

$$(A_n\psi_x)(x) = x^2(A_n f_0 - f_0)(x) - 2x(A_n f_1 - f_1)(x) + (A_n f_2 - f_2)(x).$$

因此, 当 $n \geqslant 0$ 时

$$\sup_{0\leqslant x\leqslant 1} |(A_n\psi_x)(x)| \leqslant \|A_n f_0 - f_0\| + 2\|A_n f_1 - f_1\| + \|A_n f_2 - f_2\|.$$

从而 $\lim_{n\to\infty}(\sup_{0\leqslant x\leqslant 1} |(A_n\psi_x)(x)|) = 0$, 由 Korovkin 定理即得本定理结论. □

下一定理提供了满足定理 2.13-1 假设的从 $\mathcal{C}[0,1]$ 到 $\mathcal{C}[0,1]$ 的一列线性算子 (记作 $B_n, n \geqslant 0$) 的重要例子.

定理 2.13-2 (Bernstein 定理[22]**)** 设 Bernstein 算子 $B_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1], n \geqslant 1$ 定义为对每个函数 $f : \mathcal{C}[0,1] \to \mathbb{R}$,

$$(B_n f)(x) = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}, \quad 0 \leqslant x \leqslant 1.$$

则对每个 $f \in \mathcal{C}[0,1]$,

$$\lim_{n\to\infty} \|f - B_n f\| = 0.$$

[22] S. N. Bernstein [1912]: Démonstration du théorème de Weierstrass fondée sur le calcul de probabilités. Communications of the Kharkov Mathematical Society 13. 1–2.

证明 定义 Bernstein 多项式的算子 $B_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ 显然是线性的, 且保持非负性, 由简单的计算 (习题 2.13-1) 可知, 对于 $p = 0, 1, 2$,

$$\lim_{n\to\infty} \|B_n f_p - f_p\| = 0,$$

这里 f_p 的定义见定理 2.13-1, 于是这个定理的条件满足, 从而即得结论. □

在定理 2.13-2 中定义的函数 $B_n f, n \geqslant 0$ 称为 f 的 n 次 Bernstein 多项式[23] 显然, $B_n f$ 的次数 $\leqslant n$.

注 (1) 因为对于 $f \in \mathcal{C}[0,1], \|B_n f\| \leqslant \|f\| \sup_{0\leqslant x\leqslant 1}(B_n f_0)(x) = \|f\|, (B_n f_0)(x) = 1, 0 \leqslant x \leqslant 1$, 所以定义 Bernstein 多项式的算子 $B_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ 是连续的, 而且 $\|B_n\| \leqslant 1$. 实际上, 因为 $\|B_n f_0\| = \|f_0\| = 1$, 故而 $\|B_n\| = 1$.

(2) 对 $\|f - B_n f\|$ 收敛于 0 的速度的估计, 可在 $f \in \mathcal{C}^2[0,1]$ 的附加假设下得到, 见习题 2.13-3.

作为一个直接推论, Bernstein 定理提供了分析中最基本的定理之一的一个构造证明. 用

$$\mathcal{P}[0,1], \text{ 或相应地 } \mathcal{P}([0,1];\mathbb{C})$$

表示带有实系数或相应的复系数的 $[0,1]$ 上实变量多项式全体组成的实或相应的复向量空间.

定理 2.13-3 (Weierstrass 多项式逼近定理[24]) 空间 $\mathcal{P}[0,1]$ 在赋以 sup 范数的空间 $\mathcal{C}[0,1]$ 中稠密. 类似地, 空间 $\mathcal{P}([0,1];\mathbb{C})$ 在赋以 sup 范数的空间 $\mathcal{C}([0,1];\mathbb{C})$ 中稠密.

证明 对任意给定的 $f \in \mathcal{C}[0,1]$, 由 Bernstein 定理可知由 Bernstein 多项式构成的序列 $(B_n f)_{n=1}^{\infty}$ 当 $n \to \infty$ 时一致收敛于 f, 因此, $\mathcal{P}[0,1]$ 在 $\mathcal{C}[0,1]$ 中稠密.

复数的情况可把同样的讨论用于 $\mathcal{C}([0,1];\mathbb{C})$ 中函数的实部与复部即可. □

Weierstrass 定理有一个有趣的推论.

定理 2.13-4 赋以 sup 范数的空间 $\mathcal{C}[0,1]$ 和 $\mathcal{C}([0,1];\mathbb{C})$ 均是可分的.

证明 设函数 $f \in \mathcal{C}[0,1]$, 给定 $\varepsilon > 0$. 由 Weierstrass 逼近定理, 存在实系数 c_k 的多项式 $p : x \in [0,1] \to \sum_{k=0}^{n} c_k x^k$, 使得

$$\|f - p\| < \frac{\varepsilon}{2}.$$

---

[23] 对 Bernstein 多项式的扩充讨论可见:

G. G. Lorentz [1986]: Bernstein Polynomials. Chelsea, New York.

R. Devore, G. G. Lorentz [1993]: Constructive Approximation. Springer, Heidelberg.

[24] K. Weierstrass [1885]: Über die analytische Darstelbarkeit segenannter wilkürlicher Funktionen einer reellen Veränderlichen. Sitzungsberichte der Akademie zu Berlin, 633–639 and 789–805.

因为 $\overline{\mathbb{Q}} = \mathbb{R}$, 存在 $r_k \in \mathbb{Q}$, 使得 $|c_k - r_k| \leqslant \frac{\varepsilon}{2(n+1)}, 0 \leqslant k \leqslant n$. 令 $q(x) = \sum_{k=0}^{n} r_k x^k, 0 \leqslant x \leqslant 1$, 则有

$$\|p - q\| \leqslant \sup_{0 \leqslant x \leqslant 1} \left(\sum_{k=0}^{n} |c_k - r_k| x^k \right) \leqslant \frac{\varepsilon}{2},$$

所以 $\|f - q\| \leqslant \varepsilon$. 再注意到有理系数的多项式组成的集合是可数集即可.

复的情况可以类似处理. □

在结束空间 $\mathcal{C}[0,1]$ 上多项式逼近的分析时, 我们再介绍一下十分特殊却又值得注意的结果. 为此, 先需要一个一般性的定义: 设 A 是赋范向量空间 X 的一个子集, 如果 $\overline{\operatorname{Span} A} = X$, 即 A 中元素的有限线性组合全体组成的 X 的子空间在 X 中稠密, 则称 A 是完全的.

Weierstrass 多项式逼近定理说明集合 $A = \cup_{n=0}^{\infty}\{p_n\}$ 在 $\mathcal{C}[0,1]$ 中是完全的, 这里 $p_n(x) = x^n, 0 \leqslant x \leqslant 1$. 自然会提出一个问题: 是否有由 x 的幂组成的 $\mathcal{C}[0,1]$ 的其他子集在 $\mathcal{C}[0,1]$ 中完全? 下面的定理对这个问题给了一个漂亮的回答.

定理 2.13-5 (Müntz 定理[25]) 设 $\alpha_0 = 0 < \alpha_1 < \alpha_2 < \cdots < \alpha_n < \cdots$ 满足 $\lim_{n\to\infty} \alpha_n = \infty$ (数 $\alpha_n, n \geqslant 1$, 不必是整数), 令 $e_n(x) := x^{\alpha_n}, 0 \leqslant x \leqslant 1$, 则集合 $A = \cup_{n=0}^{\infty}\{e_n\}$ 在 $\mathcal{C}[0,1]$ 中完全, 当且仅当级数 $\sum_{n=1}^{\infty} \frac{1}{\alpha_n}$ 发散.

习题

2.13-1 证明对定理 2.13.1 中的函数 $f_p, p = 0, 1, 2$, 当 $n \geqslant 2$ 时的 Bernstein 多项式为

$$(B_n f_0)(x) = 1, \ (B_n f_1)(x) = x,$$
$$(B_n f_2)(x) = x^2 + \frac{x - x^2}{n}, \ 0 \leqslant x \leqslant 1,$$

由此证明对 $p = 0, 1, 2, \lim_{n\to\infty} \|B_n f_p - f_p\| = 0$ (定理 2.13-2 证明中用到的性质).

2.13-2 设函数 $f \in \mathcal{C}[0,1]$ 定义为 $f(x) = \sqrt{x}, 0 \leqslant x \leqslant 1$, 多项式 $p_n \in \mathcal{P}[0,1], n \geqslant 0$, 递推定义为

$$p_0(x) = 0, \ p_n(x) = p_{n-1}(x) + \frac{1}{2}(x - [p_{n-1}(x)]^2), \quad n \geqslant 1, \quad 0 \leqslant x \leqslant 1.$$

证明 $\lim_{n\to\infty} \|p_n - f\| = 0$ (这个例子提供了对这个特殊的函数 f 可构造出一致收敛于它的具体的多项式, 而 Bernstein 多项式是另一种例子).

提示: 对序列 $(p_n)_{n=0}^{\infty}$ 用 Dini 定理 (习题 2.3-1).

2.13-3 证明: 对函数 $f \in \mathcal{C}^2[0,1]$, 其 Bernstein 多项式 $B_n f, n \geqslant 0$, 满足

$$f(x) - B_n f(x) = \frac{1}{n} \frac{x(x-1)}{2} f''(x) + \frac{1}{n} \varepsilon(n, x), \quad 0 \leqslant x \leqslant 1,$$

[25] C. Müntz [1914]: Über den Approximationssatz von Weierstrass. H. A. Schwarz Festschrift, pp. 303–312, Mathematische Abhandlungen, Springer, Berlin.

证明也可见 Goffman & Pedrick [1965] 第 4 章第 4.6 节或 Cheney [1966] 第 6 章第 2 节.

这里 $\lim_{n\to\infty}(\sup_{0\leqslant x\leqslant 1}|\varepsilon(n,x)|)=0$. 这个性质构成了 Voronovskaja 定理[26].

2.14 Korovkin 定理应用于三角多项式逼近; Fejér 定理

Korovkin 定理的第二个应用是在空间

$$\mathcal{C}_{\rm per}[0,2\pi]$$

上, 这个空间由所有以 2π 为周期的连续函数 $g:[0,2\pi]\to\mathbb{R}$ 组成, 其上赋以 sup 范数, 即 $\|g\|:=\sup_{0\leqslant\theta\leqslant 2\pi}|g(\theta)|$. 注意它与定理 2.13-1 惊人地类似.

定理 2.14-1[27] 设线性算子 $A_n:\mathcal{C}_{\rm per}[0,2\pi]\to\mathcal{C}_{\rm per}[0,2\pi]$ 的序列 $(A_n)_{n=0}^{\infty}$ 具有以下两个性质: 第一, 每个 $n\geqslant 0$, A_n 具有保持非负性, 即对 $g\in\mathcal{C}_{\rm per}[0,2\pi]$, 且当 $g(\theta)\geqslant 0$, $0\leqslant\theta\leqslant 2\pi$ 时, 对 $0\leqslant\theta\leqslant 2\pi$, 均有

$$(A_ng)(\theta)\geqslant 0,$$

第二, 对 $p=0,1,2$,

$$\lim_{n\to\infty}\|g_p-A_ng_p\|=0,$$

其中当 $p=0,1,2$ 时 $g_p\in\mathcal{C}_{\rm per}[0,2\pi]$ 定义为

$$g_0(\theta)=1,\quad g_1(\theta)=\cos\theta,\quad g_2(\theta)=\sin\theta,\quad 0\leqslant\theta\leqslant 2\pi.$$

则对每个 $g\in\mathcal{C}_{\rm per}[0,2\pi]$, 均有

$$\lim_{n\to\infty}\|g-A_ng\|=0.$$

证明 在集合

$$K:=\{x=(x_1,x_2)\in\mathbb{R}^2;x^2+x_2^2=1\}$$

上赋以 $\mathbb{R}^2$ 上 Euclid 范数导出的距离 d, 对任意给定的函数 $g\in\mathcal{C}_{\rm per}[0,2\pi]$, 定义函数 $g^{\#}:K\to\mathbb{R}$ 为

$$g^{\#}(x):=g(\theta),\quad x=(\cos\theta,\sin\theta),\quad 0\leqslant\theta<2\pi.$$

---

[26] E. V. Voronovskaja [1932]: Determination de la forme asymptotique de l'approximation des fonctions par les polynômes de M. Bernstein. Doklady Akademii Nauk SSSR 4, 79–85.

这个结果立即被推广到函数 $f\in\mathcal{C}^{2p}[0,1], p\geqslant 1$, 见:

S. N. Bernstein [1932]: Complément à l'article de E. Voronovskaya: "Détermination de la forme asymptotique de l'approximation des fonctions par les polynômes de M. Bernstein." Doklady Akademii Nauk SSSR 4, 86–92.

[27] P. P. Korovkin [1959]: Linear Operators and Approximation Theory. Fitzmatgiz, Moscow (俄语) [英译本, Hindustan Publishing Corporation, Delhi, 1960].

因为当 $|\theta-\varphi|\leqslant\frac{\pi}{2}$ 时

$$\frac{1}{\sqrt{2}}|\theta-\varphi|\leqslant d((\cos\theta,\sin\theta),\ (\cos\varphi,\sin\varphi))\leqslant|\theta-\varphi|,$$

所以 $g^{\#}$ 属于 $\mathcal{C}(K)$ (连续性是局部性质), 又因为函数 g 以 2π 为周期, 所以当 $\theta\in[0,2\pi]$ 且趋于 2π 时, $g(\theta)$ 趋于 $g(0)$. 显然, 这样定义的映射

$$g\in\mathcal{C}_{\text{per}}[0,2\pi]\to g^{\#}\in\mathcal{C}(K)$$

是一个双射.

下面我们证明 Korovkin 定理 (定理 2.12-1) 可以应用于空间 $\mathcal{C}(K)$, 其上赋以 sup 范数, 也记作 $\|\cdot\|$, 相应的函数 $\varphi\in\mathcal{C}[0,\infty[$ 定义为当 $t\geqslant0$ 时 $\varphi(t)=t^2$(如同定理 2.13-1 证明中出现的), 相应的算子 $A_n^{\#}:\mathcal{C}(K)\to\mathcal{C}(K),n\geqslant0$ 定义为

$$A_n^{\#}g^{\#}=(A_ng)^{\#},\quad g\in\mathcal{C}_{\text{per}}[0,2\pi].$$

为此, 首先注意到 (K,d) 为紧距离空间 (作为 $\mathbb{R}^2$ 的有界闭子集), 线性算子 $A_n^{\#}$ 也具有保持非负性, 而且对于 $f_0(x):=1,x\in K$, 因为 $f_0=g_0^{\#}$, 故而 $\lim_{n\to\infty}\|A_n^{\#}f_0-f_0\|=0$. 现在只要验证

$$\lim_{n\to\infty}(\sup_{x\in K}|(A_n^{\#}\psi_x^{\#})(x)|)=0,$$

这里, 对一切 $x=(\cos\theta,\sin\theta)\in K$, 函数 $\psi_x^{\#}\in\mathcal{C}(K)$ 定义为

$$\begin{aligned}\psi_x^{\#}(y)&=\varphi(d(x,y))=|d(x,y)|^2=4\sin^2\left(\frac{\theta-\varphi}{2}\right)\\&=2g_0(\varphi)-2\cos\theta g_1(\varphi)-2\sin\theta g_2(\varphi),\end{aligned}$$

其中 $y=(\cos\varphi,\sin\varphi)\in K$, 这又等价于 $\psi_x^{\#}=2g_0^{\#}-2\cos\theta g_1^{\#}-2\sin\theta g_2^{\#}$. 于是

$$A_n^{\#}\psi_x^{\#}=2(A_ng_0-\cos\theta(A_ng_1)-\sin\theta(A_ng_2))^{\#},$$

由此, 对一切 $x=(\cos\theta,\sin\theta)\in K$,

$$\begin{aligned}(A_n^{\#}\psi_x^{\#})(x)&=2(A_ng_0-\cos\theta(A_ng_1)-\sin\theta(A_ng_2))^{\#}(x)\\&=2A_ng_0(\theta)-2\cos\theta(A_ng_1)(\theta)-2\sin\theta(A_ng_2)(\theta).\end{aligned}$$

因为当 $0\leqslant\theta\leqslant2\pi$ 时 $g_0(\theta)-\cos\theta g_1(\theta)-\sin\theta g_2(\theta)=0$, 上述最后的关系式也可表为对一切 $x=(\cos\theta,\sin\theta)\in K$, 有

$$\begin{aligned}(A_n^{\#}\psi_x^{\#})(x)=2(A_ng_0-g_0)(\theta)-2\cos\theta(A_ng_1-g_1)(\theta)\\-2\sin\theta(A_ng_2-g_2)(\theta).\end{aligned}$$

由此即得

$$\sup_{x\in K}|(A_n^{\#}\psi_x^{\#})(x)|\leqslant2(\|A_ng_0-g_0\|+\|A_ng_1-g_1\|+\|A_ng_2-g_2\|),$$

故而 $\lim_{n\to\infty}\sup_{x\in K}|(A_n^{\#}\psi_x^{\#})(x)|=0$. 由 Korovkin 定理即得要证的结论. □

下面的定理将要给出由 $\mathcal{C}_{\text{per}}[0,2\pi]$ 到 $\mathcal{C}_{\text{per}}[0,2\pi]$ 的线性算子序列的一个重要的例子 (当 $n \geqslant 0$ 时这些算子记作 F_n), 它们满足定理 2.14-1 的假设.

首先, 逐步给出一些定义. 对任何整数 $n \geqslant 0$, 用

$$\mathcal{Q}_n[0,2\pi]$$

表示所有次数 $\leqslant n$ 的以 2π 为周期的实三角多项式构成的空间, 即其元素为 $\mathcal{C}_{\text{per}}[0,2\pi]$ 中形如

$$\theta \in [0,2\pi] \to \sum_{k=0}^{n} c_k \cos k\theta + \sum_{k=1}^{n} d_k \sin k\theta$$

的函数, 这里系数 c_k, d_k 均为实数. 用

$$\mathcal{Q}[0,2\pi] := \bigcup_{n=0}^{\infty} \mathcal{Q}_n[0,2\pi] \subset \mathcal{C}_{\text{per}}[0,2\pi]$$

表示以 2π 为周期的所有实三角多项式组成的空间. 显然, $\dim \mathcal{Q}_n[0,2\pi] = 2n+1$. 下面定理中的 $S_n g$ 和 $F_{n+1} g$ 提供了次数 $\leqslant n$ 的这类三角多项式的例子. 如同我们将在后面看到的 (定理 4.9-2), 函数 $S_n g$ 的名称来自它本身就是函数 g 的 Fourier 级数的第 n 个部分和.

定理 2.14-2 (Fejér 定理[28]) 对任何整数 $n \geqslant 0$, Fourier 部分和算子 $S_n : g \in \mathcal{C}_{\text{per}}[0,2\pi] \to S_n g \in \mathcal{C}_{\text{per}}[0,2\pi]$ 定义为

$$(S_0 g)(\theta) := \frac{a_0}{2},$$
$$(S_n g)(\theta) := \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos k\theta + b_k \sin k\theta),$$

其中 $n \geqslant 1, 0 \leqslant \theta \leqslant 2\pi$,

$$a_k := \frac{1}{\pi} \int_0^{2\pi} g(\varphi) \cos k\varphi d\varphi, \quad k \geqslant 0,$$
$$b_k := \frac{1}{\pi} \int_0^{2\pi} g(\varphi) \sin k\varphi d\varphi, \quad k \geqslant 1.$$

对任何整数 $n \geqslant 1$, Fejér 算子 $F_n : \mathcal{C}_{\text{per}}[0,2\pi] \to \mathcal{C}_{\text{per}}[0,2\pi]$ 定义为

$$F_n : g \in \mathcal{C}_{\text{per}}[0,2\pi] \to F_n g := \frac{1}{n}(S_0 g + S_1 g + \cdots + S_{n-1} g).$$

则对任何 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$,

$$\lim_{n\to\infty} \|g - F_n g\| = 0.$$

[28] L. Fejér [1900]: Sur les fonctions bornées et intégrables. Comptes Rendus de l'Académie des Sciences, Paris 131, 984–987.

证明 Fejér 算子 F_n 显然是线性的. 对任何 $n \geqslant 1$, 由直接计算 (习题 2.14-1) 可知

$$F_n g(\theta) = \frac{1}{2n\pi}\int_0^{2\pi} g(\theta+\varphi)\left(\frac{\sin\frac{n\varphi}{2}}{\sin\frac{\varphi}{2}}\right)^2 d\varphi, \quad 0 \leqslant \theta \leqslant 2\pi,$$
$$\lim_{n\to\infty}\|F_n g_p - g_p\| = 0, \quad p = 0,1,2,$$

这里函数 $g_p, p=0,1,2,$ 的定义见定理 2.14-1. 由上面第一个式子易见算子 F_n 具有保持非负性, 从而满足定理 2.14-1 的所有假设. □

注 Fejér 算子 $F_n : \mathcal{C}_{\text{per}}[0,2\pi] \to \mathcal{C}_{\text{per}}[0,2\pi], n \geqslant 1$, 是连续的, 这是因为

$$\|F_n g\| \leqslant \|g\|\left(\frac{1}{2n\pi}\int_0^{2\pi}\left(\frac{\sin\frac{n\varphi}{2}}{\sin\frac{\varphi}{2}}\right)^2 d\varphi\right),$$
$$F_n g_0(\theta) = \frac{1}{2n\pi}\int_0^{2\pi}\left(\frac{\sin\frac{n\varphi}{2}}{\sin\frac{\varphi}{2}}\right)^2 d\varphi = 1, \quad 0 \leqslant \theta \leqslant 2\pi$$

(习题 2.14-1). 所以 $\|F_n\| \leqslant 1$. 实际上, 因为 $\|F_n g_0\| = \|g_0\| = 1$, 故而 $\|F_n\| = 1$.

对 $n \geqslant 0$, 函数 $F_n g$ 称为 g 的 Fejér *三角多项式*.

上面的定理说明, 对于 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$, 当 $n \to \infty$ 时, 其 Fejér 三角多项式 $F_n g$ 一致收敛于 g. 但是, 对第 n 个 Fourier 部分和 $S_n g$ 而言, 当 $n \to \infty$ 时 $S_n g$ 甚至未必点态收敛于 g[29], 更不用说一致收敛 (除非 g 满足附加的条件, 见习题 2.14-2). 下面我们将会看到 (4.9 节), 可以证明 $\lim_{n\to\infty}\|S_n g - g\|_{L^2(0,2\pi)} = 0$, 这里 $\|\cdot\|_{L^2(0,2\pi)}$ 表示空间 $L^2(0,2\pi)$ 上的范数 (2.5 节). 实际上这个关系不仅对任意 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$ 成立, 而且对任何 $g \in L^2(0,2\pi)$ 成立.

注 算子 F_n 是算子 $S_n : \mathcal{C}_{\text{per}}[0,2\pi] \to \mathcal{C}_{\text{per}}[0,2\pi]$ 的 Cesàro 平均[30] $F_n := \frac{1}{n}(S_0 + S_1 + \cdots + S_{n-1})$. 如同这里所见, "平均过程" 经常改善了收敛性质.

Fejér 定理直接提供了另一个基本结果的构造性证明, 其等价于 Weierstrass 逼近定理 (定理 2.13-3), 它构造的是三角多项式, 将这个结果应用于实空间 $\mathcal{C}_{\text{per}}[0,2\pi]$ (定义如前) 或由所有以 2π 为周期的连续函数 $g : [0,2\pi] \to \mathbb{C}$ 组成的复空间

$$\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C}),$$

其范数 $\|\cdot\|$ 定义为 $\|g\| := \sup_{0\leqslant\theta\leqslant 2\pi}|g(\theta)|$. 对每个整数 $n \geqslant 0$, 用

$$\mathcal{Q}_n([0,2\pi];\mathbb{C})$$

---

[29] 第一个连续的周期函数的 Fourier 级数在一点发散的例子见:

P. Du Bois-Raymond [1876]: Untersuchungen über die Convergenz und Divergenz der Fourierschen Darstellungsformeln. Abhandlungen der Mathematisch-Physikalischen Klasse der Königlich Bayerischen Akademie der Wissenschaften 12, 1–103.

[30] 这个命名是为了纪念 Ernesto Cesàro (1859—1906).

表示由所有次数 $\leqslant n$ 的以 2π 为周期的复三角多项式, 即 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 中形如

$$\theta \in [0,2\pi] \to \sum_{k=-n}^{n} c_k e^{ik\theta}$$

的函数组成的空间, 其中 c_k 为复系数. 最后, 用

$$\mathcal{Q}([0,2\pi];\mathbb{C}) := \bigcup_{n=0}^{\infty} \mathcal{Q}_n([0,2\pi];\mathbb{C}) \subset \mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$$

表示由所有以 2π 为周期的复三角多项式组成的空间.

定理 2.14-3 (Weierstrass 三角多项式逼近定理) 由所有以 2π 为周期的实三角多项式构成的空间 $\mathcal{Q}[0,2\pi]$ 在空间 $\mathcal{C}_{\text{per}}[0,2\pi]$ 中稠密.

类似地, 由所有以 2π 为周期的复三角多项式构成的空间 $\mathcal{Q}([0,2\pi];\mathbb{C})$ 在空间 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 中稠密.

证明 对任意给定的函数 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$, 当 $n \to \infty$ 时, 序列 $(F_n g)_{n=1}^{\infty}$ 一致收敛于 g, 其中 F_n 表示 Fejér 算子 (定理 2.14-2) 因此 $\mathcal{Q}[0,2\pi]$ 在 $\mathcal{C}_{\text{per}}[0,2\pi]$ 中稠密.

将同样的讨论用于复值函数 $g \in \mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 的实部与虚部, 即知 g 可以用特殊的形为

$$\theta \in [0,2\pi] \to \sum_{k=0}^{n} a_k \cos k\theta + \sum_{k=1}^{\infty} b_k \sin k\theta$$

的三角多项式一致逼近, 其中 a_k, b_k 为复系数. 而这样为三角多项式又可立即改写为空间 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 中的复三角多项式, 即形为

$$\theta \in [0,2\pi] \to \sum_{k=-n}^{n} c_k e^{ik\theta},$$

其中 c_k 为复系数 (为此, 可令 $c_0 := a_0$, 当 $1 \leqslant k \leqslant n$ 时, $c_k := \frac{1}{2}(a_k - ib_k)$, 当 $-n \leqslant k \leqslant -1$ 时, $c_k := \frac{1}{2}(a_{-k} + ib_{-k})$). □

我们将在后面 (定理 2.15-4) 看到, 在复的情况下, 可以由 Weierstrass 逼近定理的更强的形式得到同样的结论. 这个更强的形式便是下一节的课题.

习题

2.14-1 (1) 给定函数 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$, 证明对任何 $n \geqslant 0$, g 的第 n 个 Fourier 部分和 (定理 2.14-2) 也可用

$$(S_n g)(\theta) = \frac{1}{2\pi} \int_0^{2\pi} g(\theta + \varphi) \frac{\sin(n+\frac{1}{2})\varphi}{\sin \frac{1}{2}\varphi} d\varphi$$

给出. 出现在这个公式中的函数 $\varphi \in [0,2\pi] \to \frac{1}{2\pi} \frac{\sin(n+\frac{1}{2})}{\sin \frac{1}{2}\varphi}\varphi$ 称为 Dirichlet 核 (当 $\varphi = 0$ 时定义为 $n + \frac{1}{2}$).

(2) 对任何 $n \geqslant 1$, 用 F_n 表示 Fejér 算子 (定理 2.14-2), 证明函数 $F_n g$ 也可表示为

$$(F_n g)(\theta) = \frac{1}{2n\pi}\int_0^{2\pi} g(\theta+\varphi)\left(\frac{\sin\frac{n\varphi}{2}}{\sin\frac{\varphi}{2}}\right)^2 d\varphi.$$

公式中出现的函数 $\varphi \in [0,2\pi] \to \left(\frac{\sin\frac{n\varphi}{2}}{\sin\frac{\varphi}{2}}\right)^2$ 称为 Fejér 核 (自然地, 当 $\varphi = 0$ 时定义为 n^2).

(3) 设对 $p = 0,1,2$, 函数 g_p 由定理 2.14-1 定义, 证明对任何 $n \geqslant 1$,

$$(F_n g_0)(\theta) = 1, \quad (F_n g_1)(\theta) = \frac{n-1}{n}\cos\theta,$$
$$(F_n g_2)(\theta) = \frac{n-1}{n}\sin\theta, \quad 0 \leqslant \theta \leqslant 2\pi.$$

由此证明 $\lim_{n\to\infty}\|F_n g_p - g_p\| = 0$ (定理 2.14-2 证明中用到的一条性质).

2.14-2 设 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$ 在点 $\theta_0 \in [0,2\pi]$ 处可微, 证明 $\lim_{n\to\infty}(S_n g)(\theta_0) = g(\theta_0)$, 这里 $S_n g$ 表示 g 的第 n 个 Fourier 部分和 (定理 2.14-2).

提示: 应用习题 2.14-1(1).

2.15 Stone-Weierstrass 定理

一个代数是指在 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的一个向量空间, 其上还附有称为乘法的映射

$$(x,y) \in X \times X \to xy \in X,$$

满足下列性质: 对一切 $x,y,z \in X$ 和 $\alpha,\beta \in \mathbb{K}$,

$$(xy)z = x(yz),$$
$$x(y+z) = xy + xz, \quad (x+y)z = xz + yz,$$
$$(\alpha x)(\beta y) = (\alpha\beta)(xy).$$

当 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 时, 相应地称 X 为实代数或复代数.

代数 X 的子代数是指 X 的一个子空间, 要求它本身也是一个代数. 例如, 空间 $\mathcal{C}[0,1]$ 是一个实代数, 其子空间 $\mathcal{P}[0,1]$ 是 $\mathcal{C}[0,1]$ 的一个子代数. Weierstrass 逼近定理 (定理 2.13-3) 是说子代数 $\mathcal{P}[0,1]$ 在代数 $\mathcal{C}[0,1]$ 中稠密 (关于 sup 范数).

在第 2.13 节中, Weierstrass 定理是把 Korovkin 定理的一个推论应用于 Bernstein 多项式而证得的, 它还有一个完全不同的证明, 这个证明只是利用 $\mathcal{P}[0,1]$ 是代数 $\mathcal{C}[0,1]$ 的具有下列两个特殊性质的子代数的事实, 其一是它包含常数, 其二它分离 $\mathcal{C}[0,1]$ 中的点, 即任意给定两个不同的点 $\xi,\eta \in [0,1]$; 存在函数 $g \in \mathcal{P}[0,1]$, 使得 $g(\xi) \neq g(\eta)$ (例如, 定义为 $g(x) := x, 0 \leqslant x \leqslant 1$).

值得注意的是对任意给定的紧距离空间 K 和实空间 $\mathcal{C}(K)$ 的任何子代数, 只要满足同样简单的假设, 就有同样的稠密性质. 这就是下一个定理的本质所在. 这个定理是泛函分析最重要的结果之一.

定理 2.15-1 (Stone-Weierstrass 定理[31]) 设 K 为紧距离空间, $\mathcal{A}$ 为 (实) 空间 $\mathcal{C}(K)$ 的子代数, 具有下述两个性质:

(1) 常数函数属于 $\mathcal{A}$.

(2) 对任意给定的两个不同的点 $\xi, \eta \in K$, 存在函数 $g = g(\xi, \eta) \in \mathcal{A}$, 使得 $g(\xi) \neq g(\eta)$.

则 $\mathcal{A}$ 在 $\mathcal{C}(K)$ 中稠密.

证明 (i) $\mathcal{A}$ 的闭包 $\overline{\mathcal{A}}$ 也是 $\mathcal{C}(K)$ 的一个子代数.

这是因为加法与数乘是连续的 (定理 2.2-5), 乘法运算作为 $\mathcal{C}(K) \times \mathcal{C}(K) \to \mathcal{C}(K)$ 的映射也是连续的. 这里, 乘法的连续性可以由等式 $\tilde{f}\tilde{g} - fg = (\tilde{f} - f)g + (\tilde{g} - g)f + (\tilde{f} - f)(\tilde{g} - g)$ 和不等式 $\|fg\| \leqslant \|f\|\,\|g\|$ 得到, 这里还利用了 $\mathcal{C}(K)$ 中乘法是可交换的性质.

(ii) 如果 $f \in \overline{\mathcal{A}}$, 则 $|f| \in \overline{\mathcal{A}}$.

首先, 注意到当 $f \in \mathcal{C}(K)$ 时 $|f| \in \mathcal{C}(K)$ (因为对任意的 $x, y \in K, \big|\,|f(x)| - |f(y)|\,\big| \leqslant |f(x) - f(y)|$).

其次, 给定 $f \in \overline{\mathcal{A}}$ 和 $\varepsilon > 0$. 无需依赖 Bernstein 多项式 (定理 2.13-2) 或 Weierstrass 多项式逼近定理 (定理 2.13-2), 显然 (习题 2.15-3) 存在多项式 $p \in \mathcal{P}$, 使得

$$\sup_{-\|f\| \leqslant t \leqslant \|f\|} \big|\,|t| - p(t)\big| \leqslant \varepsilon.$$

因此,

$$\sup_{x \in K} \big|\,|f(x)| - p(f(x))\big| = \|f - p \circ f\| \leqslant \varepsilon.$$

因为 p 是多项式, 由 (i) $\overline{\mathcal{A}}$ 也是子代数, 所以函数 $p \circ f$ 也属于 $\overline{\mathcal{A}}$. 由 $\varepsilon > 0$ 的任意性, 函数 $|f| \in \mathcal{C}(K)$ 也属于 $\overline{\mathcal{A}}$.

(iii) 如果 $f, g \in \overline{\mathcal{A}}$, 则 $\max\{f, g\} \in \overline{\mathcal{A}}, \min\{f, g\} \in \overline{\mathcal{A}}$.

为此, 只要由关系式

$$\max\{f, g\} = \frac{1}{2}(f + g + |f - g|),$$
$$\min\{f, g\} = \frac{1}{2}(f + g - |f - g|),$$

并利用 (ii) 即得.

(iv) 任意给定点 $\xi, \eta \in K$ 和 $\alpha, \beta \in \mathbb{R}$, 存在函数 $g \in \mathcal{A}$, 使得 $g(\xi) = \alpha, g(\eta) = \beta$.

由假设, 存在函数 $g_0 \in \mathcal{A}$, 使得 $g_0(\xi) \neq g_0(\eta)$. 定义函数 $g \in \mathcal{C}(K)$ 为

$$g(x) := \frac{\alpha g_0(\eta) - \beta g_0(\xi)}{g_0(\eta) - g_0(\xi)} + \frac{\beta - \alpha}{g_0(\eta) - g_0(\xi)} g_0(x), \quad x \in K,$$

[31] M. H. Stone [1948]: The generalized Weierstrass approximation theorem. Mathematics Magazine 21, 167–183 与 237–254.

于是 g 属于 $\mathcal{A}$ (由假设, 常数函数属于 $\mathcal{A}$, 又 $g_0 \in \mathcal{A}$), 且满足 $g(\xi) = \alpha, g(\eta) = \beta$.

(v) 给定函数 $f \in \mathcal{C}(K)$ 和 $\varepsilon > 0$, 存在函数 $g \in \overline{\mathcal{A}}$, 使得 $\|f - g\| \leqslant \varepsilon$.

由 (iv), 对任意给定的 $\xi, \eta \in K$ 存在函数 $g(\xi,\eta) \in \mathcal{A}$, 使得

$$g(\xi,\eta)(\xi) = f(\xi), \quad g(\xi,\eta)(\eta) = f(\eta).$$

每个集合

$$U(\xi,\eta) := \{x \in K; g(\xi,\eta)(x) < f(x) + \varepsilon\}$$

是 K 中开集 (因为函数 $g(\xi,\eta)$ 和 f 均连续), 对所有的 $\eta \in K$, 因为当 $\xi, \eta \in K$ 时, $\xi \in U(\xi,\eta)$, 所以 $K = \cup_{\xi \in K} U(\xi,\eta)$. 因为 K 是紧集, 故对上述的 K 的开覆盖, 存在有限子覆盖, 这样

$$K = \bigcup_{i=1}^{m(\eta)} U(\xi_i, \eta).$$

对每个 $\eta \in K$, 定义函数

$$g(\eta) := \min_{1 \leqslant i \leqslant m(\eta)} \{g(\xi_i, \eta)\},$$

由 (iii) 知, 它属于 $\overline{\mathcal{A}}$ (这就是为什么这里需要 "有限最小值"). 任意给定 $x \in K$, 存在 $i = i(x,\eta) \in \{1, 2, \cdots, m(\eta)\}$, 使得 $x \in U(\xi_i, \eta)$, 因此 $g(\xi_i,\eta)(x) < f(x) + \varepsilon$. 这样, 对所有的 $x \in K$,

$$g(\eta)(x) \leqslant g(\xi_{i(x,\eta)}, \eta)(x) < f(x) + \varepsilon.$$

每个集合

$$V(\eta) := \{x \in K; g(\eta)(x) > f(x) - \varepsilon\}$$

是 K 中的开集 (因为 $g(\eta)$ 与 f 均连续), 又对一切 $\eta \in K, \eta \in V(\eta)$, 所以 $K = \cup_{\eta \in K} V(\eta)$. 于是, 存在 K 的有限子覆盖即 K 形为

$$K = \bigcup_{j=1}^{n} V(\eta_j).$$

定义函数

$$g := \max_{1 \leqslant j \leqslant n} \{g(\eta_j)\},$$

由 (iii) 知, 它属于 $\overline{\mathcal{A}}$ (这就是为什么这里需要 "有限最大值") 给定任意的 $x \in K$, 存在 $j = j(x) \in \{1, 2, \cdots, n\}$, 使得 $x \in V(\eta_j)$, 这样, $g(\eta_j)(x) > f(x) - \varepsilon$. 因此, 一方面对一切 $x \in K$,

$$g(x) \geqslant g(\eta_{j(x)})(x) > f(x) + \varepsilon.$$

另一方面, 存在 $k = k(x) \in \{1, 2, \cdots, n\}$, 使得 $g(x) = g(\eta_{k(x)})(x)$, 于是, 对一切 $x \in K$,

$$g(x) = g(\eta_{k(x)})(x) \leqslant g(\xi_{i(x,\eta_{k(x)})}, \eta_{k(x)})(x) < f(x) + \varepsilon.$$

这就找到了函数 $g \in \overline{\mathcal{A}}$, 使得

$$\|f - g\| = \sup_{x \in K} |f(x) - g(x)| < \varepsilon.$$

由 $\varepsilon > 0$ 的任意性, 即知 f 属于 $\overline{\mathcal{A}}$ 的闭包, 而 $\overline{\mathcal{A}}$ 是闭的, 故 $\overline{\mathcal{A}}$ 的闭包即其自身. □

Stone-Weierstrass 定理的第一个推论是在实的情况下多项式逼近的经典的 Weierstrass 定理 (定理 2.13-3) 的下述推广. 另一个有趣的推论是习题 2.15-1.

定理 2.15-2 (多变量的 Weierstrass 多项式逼近定理) 设 K 为 $\mathbb{R}^n$ 的紧子集, $\mathcal{P}(K)$ 是由所有 n 元实多项式在 K 上的限制组成的空间. 则 $\mathcal{P}(K)$ 在 $\mathcal{C}(K)$ 中稠密.

证明 显然, $\mathcal{P}(K)$ 是包含常数函数的子代数. 如果 $\xi = (\xi_i)_{i=1}^n$ 和 $\eta = (\eta_i)_{i=1}^n$ 是 K 中两个不同的点, 则必存在 $i \in \{1, 2, \cdots, n\}$, 使得 $\xi_i \neq \eta_i$. 这样, 定义多项式 g 为当 $x = (x_j)_{j=1}^n$ 时 $g(x) := x_i$, 它满足 $g(\xi) \neq g(\eta)$. 由 Stone-Weierstrass 定理即得结论. □

Stone-Weierstrass 定理另一要注意的要点是它关于复值函数的下述推广. 注意, 此时需要一个额外的假设 (见下面的 (c)).

定理 2.15-3 (复 Stone-Weierstrass 定理) 设 K 为紧距离空间, $\mathcal{A}$ 为 (复) 空间 $\mathcal{C}(K; \mathbb{C})$ 的复子代数, 具有以下三个性质:

(a) 常数函数属于 $\mathcal{A}$.

(b) 任意给定两个不同的点 $\xi, \eta \in K$, 存在函数 $g \in \mathcal{A}$, 使得 $g(\xi) \neq g(\eta)$.

(c) 如果 $g \in \mathcal{A}$, 则其共轭函数 $\bar{g} \in \mathcal{A}$.

则 $\mathcal{A}$ 在 $\mathcal{C}(K; \mathbb{C})$ 中稠密.

证明 由于 (c), 任何函数 $f \in \mathcal{A}$ 的实部与虚部, 即

$$\operatorname{Re} f := \frac{1}{2}(f + \bar{f}), \quad \operatorname{Im} f := \frac{1}{2i}(f - \bar{f})$$

属于 $\mathcal{C}(K; \mathbb{C})$ 的子代数 $\mathcal{A}_{\mathbb{R}}$, 这个子代数由 $\mathcal{A}$ 中所有实值连续函数组成. 因此, 只要证明 $\mathcal{A}_{\mathbb{R}}$ 满足 (实) Stone-Weierstrass 定理 (定理 2.15-1) 的假设, 并把这个定理应用于任意函数 $f \in \mathcal{C}(K; \mathbb{C})$ 的实部与虚部即可.

由 (a), $\mathcal{A}_{\mathbb{R}}$ 显然包含实的常数函数. 任意给定两个不同的点 $\xi, \eta \in K$, 存在函数 $g \in \mathcal{A}$, 使得 $g(\xi) = 0, g(\eta) = 1$(利用 (b) 和定理 2.15-1 证明中的 (iv)). 因此, 实值函数 $\operatorname{Re} g$ 属于 $\mathcal{A}$, 且满足 $\operatorname{Re} g(\xi) = 0$, $\operatorname{Re} g(\eta) = 1$. □

复 Stone-Weierstrass 定理的一个直接推论是下述逼近论中的一个基本结果, 正如已在定理 2.14-3 中所见, 它也可从实的 Weierstrass 三角多项式逼近定理导出.

定理 2.15-4 (复三角多项式逼近定理) 设 $\mathcal{Q}([0,2\pi];\mathbb{C})$ 为所有以 2π 为周期的复三角多项式组成的空间, 即其元为 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 中形为

$$\theta \in [0,2\pi] \to \sum_{k=-n}^{n} c_k e^{ik\theta}$$

的函数 (2.14 节), 其中 c_k 为复系数. 则 $\mathcal{Q}([0,2\pi];\mathbb{C})$ 是空间 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 的子代数, 而且它在 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 中稠密.

证明 类似于定理 2.14-1 的证明, 函数 $g \in \mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 恒同于函数 $g^\# \in \mathcal{C}(K;\mathbb{C})$, 这里

$$K := \{x = (x_1, x_2) \in \mathbb{R}^2; x_1^2 + x_2^2 = 1\}.$$

给定两个不同的 $\xi, \eta \in K$, 函数 $g^\#$ 满足 $g^\#(\xi) \neq g^\#(\eta)$, 这里 $g(\theta) := e^{i\theta}, 0 \leqslant \theta \leqslant 2\pi$, 直接应用复 Stone-Weierstrass 定理 (其他条件显然满足) 即得证. □

习题

2.15-1 设 K 为紧距离空间. 利用 Stone-Weierstrass 定理, 证明 $\mathcal{C}(K)$ 是可分的[32] ($K = [0,1]$ 的特殊情况作为 Weierstrass 多项式逼近定理的推论, 已在定理 2.13-4 中导出).

2.15-2 这个习题提供了 Lebesgue 空间 $L^p(\Omega)$, $1 \leqslant p < \infty$ 可分性 (定理 2.5-4(a)) 的另一个证明. 这个证明基于多变量的 Weierstrass 多项式逼近定理 (定理 2.15-2).

(1) 设 Ω 是 $\mathbb{R}^n$ 的任一开子集. 证明: 对每个整数 $\ell \geqslant 1$, 集合

$$K_\ell := \{x \in \Omega; \text{dist}(x, \partial\Omega) \geqslant \frac{1}{\ell}\} \cap \overline{B(0,\ell)}$$

是 Ω 的紧子集, 而且对 Ω 的任何紧子集 K, 存在 $\ell = \ell(K) \geqslant 1$, 使得 $K \subset K_\ell$.

(2) 对每个整数 $\ell \geqslant 1$, 定义集合

$$\Pi_\ell := \{q : \Omega \to \mathbb{R}; q|_{K_\ell} \text{ 是有理系数的 } n \text{ 个变量的多项式, } q|_{\Omega - K_\ell} = 0\}.$$

证明: 集合 $\Pi = \cup_{\ell=1}^{\infty} \Pi_\ell$ 是 $L^p(\Omega)$, $1 \leqslant p \leqslant \infty$, 的可数无限的子集.

(3) 设对某 $1 \leqslant p < \infty$, $f \in L^p(\Omega)$, 给定 $\varepsilon > 0$. 由定理 2.5-3, 存在函数 $g = g(f,\varepsilon) \in \mathcal{C}_C(\Omega)$, 使得 $\|f - g\|_{L^p(\Omega)} \leqslant \frac{\varepsilon}{2}$. 证明存在函数 $h = h(f,\varepsilon) \in \Pi$, 使得 $\|g - h\|_{L^p(\Omega)} \leqslant \frac{\varepsilon}{2}$, 由此证得当 $1 \leqslant p < \infty$ 时 $L^p(\Omega)$ 是可分的.

(4) 类似地证明当 $1 \leqslant p < \infty$ 时, Lebesgue 空间 $L^p(\Omega;\mathbb{C})$ 是可分的.

2.15-3 利用习题 2.13-2, 证明对任何给定的 $a > 0$, 存在一列多项式 $p_n \in \mathcal{P}$, 使得 $\lim_{n\to\infty} \sup_{-a \leqslant t \leqslant a} ||t| - p_n(t)| = 0$.

---

[32] 可见 Dieudonné [1960] 定理 7.4.4.

2.16 凸集

给定向量空间 X 的两点 a 和 b, 集合

$$[a,b] := \{x \in X; x = \lambda a + (1-\lambda)b,\ 0 \leqslant \lambda \leqslant 1\},$$

称 $[a,b]$ 为以 a,b 为端点的线段或闭线段.

向量空间 X 的子集 A 若包含两点 a,b, 则包含线段 $[a,b]$, 称集合 A 为凸集. 注意, 空集和 X 中的单点集均为 X 中的凸集, 一族凸子集 $A_i \subset X$ 的交集 $\bigcap_{i\in I} A_i$ 也是凸集.

赋范向量空间中凸子集 A 的闭包 $\overline{A}$ 是凸的 (对任意给定的 $a,b \in \overline{A}$ 和任意的 $0 \leqslant \lambda \leqslant 1$, 取 $a_k, b_k \in A$, 使 $\lim_{k\to\infty} a_k = a$, $\lim_{k\to\infty} b_k = b$, 注意到对一切 k, $\lambda a_k + (1-\lambda)b_k \in A$ 即可), 类似地, (开) 球及其闭包是赋范向量空间的凸子集.

凸子集的内部是凸集 (证明见习题 2.16-2). 顺便注意, 在无限维空间中, 凸集的内部可能为空集, 见习题 2.16-7.

给定向量空间 X 的一个子集 A, X 中所有包含 A 的凸子集的交, 即包含 A 的最小凸子集, 称为 A 的凸包, 记作

$$\text{co}\, A.$$

下面的结果给出了凸包的一个常用特征. 凸包的其他一些有用的性质在习题 2.16-4 至 2.16-6 中给出.

定理 2.16-1 设 A 是向量空间 X 的子集. 则 A 的凸包就是由 A 的元素的所有凸组合构成的 X 的子集, 所谓 A 中元素的凸组合即元素 $a_i \in A$ 的有限线性组合 $\sum_{i\in I} \lambda_i a_i$ (2.1 节), 满足对一切 $i \in I$,

$$\lambda_i \geqslant 0, \quad \sum_{i\in I} \lambda_i = 1.$$

证明 (i) 设 C 为 X 的凸子集, 则所有形如

$$a = \sum_{i=1}^{n} \lambda_i a_i$$

的点属于 C, 其中当 $1 \leqslant i \leqslant n$ 时 $\lambda_i \geqslant 0$, $a_i \in C$, 且 $\sum_{i=1}^{n} \lambda_i = 1$.

假设这个性质对于 $1, 2, \cdots, n-1$ 成立 (对 $n = 1, 2$ 显然成立). 又设 $a \in X$ 如上所给. 因为 $a_n \in C$, 可设 $\lambda := \sum_{i=1}^{n-1} \lambda_i > 0$. 只要把 a 表示为

$$a = \sum_{i=1}^{n} \lambda_i a_i = \lambda \left(\sum_{i=1}^{n-1} \frac{\lambda_i}{\lambda} a_i \right) + (1-\lambda) a_n,$$

注意到 $0 < \lambda \leqslant 1$ 和由归纳假设 $\sum_{i=1}^{n-1} \frac{\lambda_i}{\lambda} a_i \in C$. 所以由 C 的凸性得 $a \in C$.

(ii) 令

$$T := \left\{ \sum_{i\in I} \lambda_i a_i; I \text{ 为有限集, 对 } i \in I,\ \lambda_i \geqslant 0,\ a_i \in A, \sum_{i\in I} \lambda_i = 1 \right\}.$$

由 (i), T 中的任何点均属于包含 A 的任意凸集. 所以, $T \subset \mathrm{co}\, A$.

(iii) (ii) 中定义的 T 是凸的. 这是因为对任何 $a = \sum_{i\in I} \lambda_i a_i \in T$ 和 $b = \sum_{j\in J} \mu_j b_j \in T$ 及 $0 \leqslant \nu \leqslant 1$, 有

$$\nu a + (1-\nu) b = \sum_{i\in I} (\nu\lambda_i) a_i + \sum_{j\in J} ((1-\nu)\mu_j) b_j,$$

而 $\nu\lambda_j \geqslant 0$, $(1-\nu)\mu_j \geqslant 0$ 且 $\sum_{i\in I}(\nu\lambda_i) + \sum_{j\in J}((1-\nu)\mu_j) = 1$.

因为 T 是凸的, 所以 $\mathrm{co}\, A \subset T$. □

当 $i \in I$ 时 $\lambda_i \geqslant 0$, $\sum_{i\in I} \lambda_i = 1$, 有限线性组合 $a = \sum_{i\in I} \lambda_i a_i$ (如定理 2.16-1 中出现的) 称为点 a_i $(i \in I)$ 的凸组合, 称点 a 为点 a_i 的带有权 λ_i 的重心.

例如, 设 $n+1$ 个点 $a_j = (a_{ij})_{i=1}^n \in \mathbb{R}^n$, $1 \leqslant j \leqslant n+1$, 它们是仿射独立的, 即并不包含于 $\mathbb{R}^n$ 中的一个超平面, 这又等价于 $(n+1)\times(n+1)$ 阵 (a_{ij}) 是可逆的, 这里 $a_{n+1,j} := 1$, $1 \leqslant j \leqslant n+1$. 这样, 由定理 2.16-1, 集合 $\cup_{j=1}^{n+1}\{a_j\}$ 的凸包可以表示为

$$T = \left\{ x \in \mathbb{R}^n; x = \sum_{j=1}^{n+1} \lambda_j a_j, \lambda_j \geqslant 0, 1 \leqslant j \leqslant n+1, \sum_{j=1}^{n+1} \lambda_j = 1 \right\},$$

称之为 n 单纯形, 点 a_j 称为其顶点 (2 单纯形为三角形, 3 单纯形为四面体). 由简单的紧性的讨论可知 T 是闭的 (这是一个一般性质的特殊情况, 见习题 2.16-5).

有限集的凸包具有下列性质.

定理 2.16-2 设 A 为赋范向量空间 X 的有限子集, 则 $\mathrm{co}\, A$ 是 X 的一个紧子集.

证明 设 $A = \bigcup_{j=1}^{m} \{x_j\} \subset X$, 则由定理 2.16-1,

$$\mathrm{co}\, A = \left\{ \sum_{j=1}^{m} \lambda_j x_j; \lambda_j \geqslant 0, 1 \leqslant j \leqslant m,\ \text{且}\ \sum_{j=1}^{m} \lambda_j = 1 \right\}.$$

任意给定无限序列 $(x^k)_{k=1}^\infty$, 其中对每个 $k \geqslant 1, x^k = \sum\limits_{j=1}^{m} \lambda_j^k x_j \in \mathrm{co}\, A$. 相应的序列 $((\lambda_1^k, \lambda_2^k, \cdots, \lambda_m^k))_{k=1}^\infty$ 在 $\mathbb{R}^m$ 中有界, 因而存在子列 $((\lambda_1^{\sigma(k)}, \lambda_2^{\sigma(k)}, \cdots, \lambda_m^{\sigma(k)}))_{k=1}^\infty$ 和元素 $(\lambda_1, \lambda_2, \cdots, \lambda_m) \in \mathbb{R}^m$, 使得对 $1 \leqslant j \leqslant m$, 当 $k \to \infty$ 时 $\lambda_j^{\sigma(k)} \to \lambda_j \geqslant 0$, 且 $\sum\limits_{j=1}^{m} \lambda_j = \lim\limits_{k\to\infty} \sum\limits_{j=1}^{m} \lambda_j^{\sigma(k)} = 1$, 从而当 $k \to \infty$ 时

$$x^{\sigma(k)} = \sum_{j=1}^{m} \lambda_j^{\sigma(k)} x_j \to \sum_{j=1}^{m} \lambda_j x_j \in \mathrm{co}\, A.$$

这就证明了 co A 是紧的. □

另一个同样重要的概念是赋范向量空间 X 的子集 A 的凸闭包.

$$\overline{\mathrm{co}}\, A,$$

其定义为 X 中所有包含 A 的闭凸子集的交, 或等价地, X 中包含 A 的最小闭凸子集.

定理 2.16-3 设 A 为赋范向量空间 X 的子集, 则 $\overline{\mathrm{co}}\, A = \overline{\mathrm{co}\, A}$.

证明 一方面, $\overline{\mathrm{co}\, A}$ 是闭集 (作为闭包), 且是凸的 (作为凸子集 co A 的闭包).

另一方面, 设 C 为 X 中包含 A 的凸闭子集. 因为任何包含 A 的凸子集包含 co A, 而 C 是凸的, 所以 C 包含 co A; 又因 C 是闭的, 所以 C 包含 $\overline{\mathrm{co}\, A}$.

因此, $\overline{\mathrm{co}\, A}$ 包含于任何包含 A 的凸闭子集之中, 且自身是包含 A 的凸闭子集, 所以 $\overline{\mathrm{co}\, A} = \overline{\mathrm{co}} A$. □

注 (1) 我们将在后面建立起完备赋范向量空间中闭凸包的一个重要性质 (定理 3.1-5).

(2) 因为 $\overline{\mathrm{co}\, A}$ 是包含 $\overline{A}$ 的闭凸集 (作为包含 A 的闭子集). 显然可得 co $\overline{A} \subset \overline{\mathrm{co}\, A}$. 这个包含关系可能是严格成立的. 例如, 在 $\mathbb{R}^2$ 中取子集 $A := \{(x_1, x_2) \in \mathbb{R}^2; x_2 \geqslant (1 + x_1^2)^{-1}\} = \overline{A}$.

习题

2.16-1 设 A 是 $\mathbb{R}^2$ 中包含原点的一个凸子集, 具有下列性质: 给定任意常数 $\alpha_1, \alpha_2 \in \mathbb{R}^2$, 满足 $|\alpha_1| + |\alpha_2| > 0$, $\mathbb{R}^2$ 中的子集 $\{(\alpha_1 t, \alpha_2 t) \in \mathbb{R}^2; t \geqslant 0\}$ (即一条从原点 0 出发的半直线) 至少包含一个不属于 A 的点. 证明集合 A 是有界的.

2.16-2 设 A 是赋范向量空间 X 的一个凸子集.

(1) 证明: 如果 $a \in \mathrm{int}\, A$, $b \in \overline{A}$, 则 $\{x \in X; x = \lambda a + (1-\lambda) b, 0 < \lambda \leqslant 1\} \subset \mathrm{int}\, A$.

(2) 证明: int A 是凸的.

(3) 证明: 如果 int $A \neq \varnothing$, 则 $\overline{A} = \overline{\mathrm{int}\, A}$ (显然, 这个性质对于 X 的任意子集 A 并不成立).

2.16-3 对任意整数 $n \geqslant 2$, 用 $\mathbb{M}^n$ 表示由所有实 $n \times n$ 阵构成的向量空间.

(1) 直接证明集合 $\mathbb{M}_+^n := \{\boldsymbol{A} \in \mathbb{M}^n; \det \boldsymbol{A} > 0\}$ 不是凸的.

(2) 证明: co $\mathbb{M}_+^n = \mathbb{M}^n$.

2.16-4 证明开集的凸包仍为开集.

2.16-5 (1) (Carathéodory 定理[33]) 设 A 为 $\mathbb{R}^n$ 的一个子集, 证明, 任意点 $x \in \mathrm{co}\, A$ 均可表示为

$$x = \sum_{i=1}^{n+1} \lambda_i a_i,$$

[33] T. Carathéodory [1907]: Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen. Rendiconti del Circolo Matematico di Palermo 32, 193–217.

其中当 $1 \leqslant i \leqslant n+1$ 时 $\lambda_i \geqslant 0$, $a_i \in A$, 且 $\sum_{i=1}^{n+1} \lambda_i = 1$, 即 x 可表示为 A 中至多 $n+1$ 个点的凸组合.

(2) 利用 (1) 证明 $\mathbb{R}^n$ 中紧子集的凸包仍是紧集.

注 我们将会看到, 性质 (2) 是 Banach 空间中更一般性质的一个特殊情况, 见定理 3.1-5.

2.16-6 (Birkhoff 定理[34]) 给定集合 $\{1, 2, \cdots, n\}$ 的一个置换 τ, 相应的 $n \times n$ 置换阵 $\boldsymbol{P}_\tau$ 定义为 $(\boldsymbol{P}_\tau)_{ij} = \delta_{i\tau(j)}$, 称 $n \times n$ 阵 (a_{ij}) 为重随机矩阵, 是指

$$a_{ij} \geqslant 0,\ 1 \leqslant i, j \leqslant n;\ \sum_{j=1}^{n} a_{ij} = 1,\ 1 \leqslant i \leqslant n;\ \sum_{i=1}^{n} a_{ij} = 1,\ 1 \leqslant j \leqslant n.$$

证明所有 $n \times n$ 置换阵组成的集合的凸包即所有 $n \times n$ 重随机矩阵组成的集合.

2.16-7 (1) 证明集合 $A_n := \{x = (x_i)_{i=1}^n \in \mathbb{R}^n; x_i \geqslant 0,\ 1 \leqslant i \leqslant n\}$ 是 $\mathbb{R}^n$ 中的凸集, 且 $\mathring{A}_n = \{x = (x_i)_{i=1}^n \in \mathbb{R}^n; x_i > 0, 1 \leqslant i \leqslant n\}$.

(2) 证明集合 $A := \{x = (x_i)_{i=1}^\infty \in \ell^2; x_i \geqslant 0, i \geqslant 1\}$ 是 ℓ^2 中的凸集, 确定其内部.

2.16-8 $n \times n$ 复矩阵 $\boldsymbol{A}$ 的数值域 $F(\boldsymbol{A})$ 定义为[35]

$$F(\boldsymbol{A}) := \{\boldsymbol{x}^* \boldsymbol{A} \boldsymbol{x}; \boldsymbol{x} \in \mathbb{C}^n, \|\boldsymbol{x}\|_2 = 1\},$$

证明 $F(\boldsymbol{A})$ 是 $\mathbb{C}$ 的凸子集.

这个 (不太容易证明的) 结果即 Toeplitz-Hausdorff 定理[36], 它也可以用以讨论无限维内积空间上的线性算子[37].

2.17 凸函数

设 X 为向量空间, A 为 X 的凸子集. 如果函数 $f: A \to \mathbb{R}$ 满足对任意两点 $a, b \in A$, 和所有的 $0 \leqslant \lambda \leqslant 1$, 均有

$$f(\lambda a + (1-\lambda) b) \leqslant \lambda f(a) + (1-\lambda) f(b),$$

则称 f 在 A 上凸. 由简单的递推即知如果 f 在 A 上凸, 则对任意的 $a_i \in A$, $1 \leqslant i \leqslant n$ 和所有的 $0 \leqslant \lambda_i \leqslant 1$, $1 \leqslant i \leqslant n$, 当 $\sum_{i=1}^n \lambda_i = 1$ 时,

---

[34] G. Birkhoff [1946]: Tres observaciones sobre el algebra lineal. Universidad Nacional de Tucumán Revista A, 5, 147–151.

[35] 矩阵的数值域是矩阵理论中一个重要的课题, 见 Horn & Johnson [1991] 第 1 章.

[36] O. Toeplitz [1918]: Das algebraische Analogon zu einem Satze von Fejér. Mathematische Zeitschrift 2, 187–197.

F. Hausdorff [1919]: Der Wertvorrat einer Bilinearform. Mathematische Zeitschrift 3, 314–316.

[37] 见 Halmos [1982] 第 22 章, 或见:

C. Davis [1971]: The Toeplitz-Hausdorff theorem explained. Canadian Mathematical Bulletin 14, 245–246.

A. Mcintosch [1978]: The Toeplitz-Hausdorff theorem and ellipticity conditions. The American Mathematical Monthly 85, 475–477.

$$f\left(\sum_{i=1}^{n}\lambda_i a_i\right) \leqslant \sum_{i=1}^{n}\lambda_i f(a_i).$$

如果对任意两个不同的点 $a, b \in A$, 和任意的 $0 < \lambda < 1$, 均有

$$f(\lambda a + (1-\lambda)b) < \lambda f(a) + (1-\lambda)f(b),$$

则称 $f: A \to \mathbb{R}$ 在 A 上严格凸.

若函数 $-g: A \to \mathbb{R}$ 是凸的, 或严格凸的, 则相应地称函数 $g: A \to \mathbb{R}$ 是凹的, 或严格凹的.

例如, 设 $f_i: A \to \mathbb{R}$, $1 \leqslant i \leqslant n$, 为凸函数, 则函数 $\max_{1\leqslant i\leqslant n}\{f_i\}$ 和 $\sum_{i=1}^{n} f_i$ 也是凸的. 如果 X 为实向量空间, 线性泛函 $\ell: X \to \mathbb{R}$ 是凸的, 但不是严格凸的. 如果 $(X; \|\cdot\|)$ 为赋范向量空间, 范数 $\|\cdot\|; X \to \mathbb{R}$ 是凸函数, 这是因为对任何 $x, y \in X$, 和一切 $0 \leqslant \lambda \leqslant 1$,

$$\|\lambda x + (1-\lambda)y\| \leqslant \lambda\|x\| + (1-\lambda)\|y\|.$$

凸性的概念可以如下自然地扩充到并非定义于凸集的函数上 (特别地, 这个扩充在多凸函数的定义中要用到, 见第 9.7 节), 设 X 为向量空间, A 为 X 的子集, 称函数 $f: A \to \mathbb{R}$ 为凸的, 是指存在凸函数 (如本节开始所理解的) $\tilde{f}: \text{co}\, A \to \mathbb{R}$, 使得 $\tilde{f}|_A = f$.

定义在有限维空间开子集上的凸函数具有值得注意的性质:

定理 2.17-1 设 Ω 为有限维空间 X 上开的凸子集, 则任何凸函数 $f: \Omega \to \mathbb{R}$ 是连续的.

证明 (i) 先证函数 f 是局部上有界的, 即对任意给定的 $a \in \Omega$, 存在一个包含在 Ω 中的邻域 B 和常数 M, 使得对一切 $x \in B, f(x) \leqslant M$.

设 $(e_i)_{i=1}^{n}$ 是空间 X 的基, 在 X 上赋以范数 $\|\cdot\|_1: x = \sum_{i=1}^{n} x_i e_i \to \sum_{i=1}^{n}|x_i|$ (因为有限维空间中所有的范数均等价, 所以这样不失一般性). 因为 Ω 是开集, 故而存在 $r > 0$, 使得

$$B := \{x \in X; \|x - a\|_1 \leqslant r\} \subset \Omega.$$

令 $a_i := a + re_i, 1 \leqslant i \leqslant n, a_i := a - re_{i-n}, n+1 \leqslant i \leqslant 2n$ (图 2.17-1). 由范数 $\|\cdot\|_1$ 的定义, 任何 $x \in B$ 均可表示为

$$x = \sum_{i=1}^{2n}\lambda_i a_i,$$

这里 $0 \leqslant \lambda_i \leqslant 1, 1 \leqslant i \leqslant 2n$ 且 $\sum_{i=1}^{2n}\lambda_i = 1$, 即 x 可表为点 a_i 的凸组合, 这里 $1 \leqslant i \leqslant 2n$. 由函数 f 的凸性可知对所有的 $x \in B$, 均有

$$f(x) \leqslant \sum_{i=1}^{2n}\lambda_i f(a_i) \leqslant M := \max_{1\leqslant i\leqslant 2n} f(a_i).$$

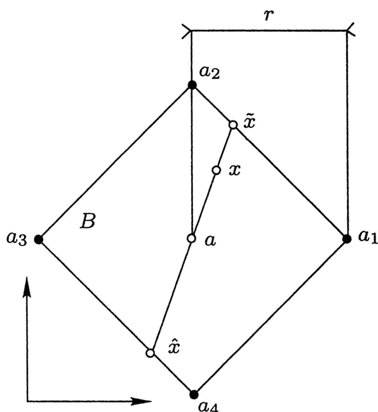


图 2.17-1 在 $\mathbb{R}^2$ 的特殊情况下, 定理 2.17-1 证明中的 $x, \tilde{x}, \hat{x}$ 和 a_i

(ii) 函数 f 是连续的.

设点 a 和闭球 B 如 (i) 中所给. 任意点 $x \in B$ 与 a 不同可以表示为 (图 2.17-1)

$$x = \lambda\widetilde{x} + (1-\lambda)a,$$

其中 $0 < \lambda := \frac{\|x-a\|_1}{r} \leqslant 1, \widetilde{x} = a + \frac{1}{\lambda}(x-a)$. 因此, $f(x) \leqslant \lambda f(\widetilde{x}) + (1-\lambda)f(a)$, 这样, 对任何 $x \in B$,

$$f(x) - f(a) \leqslant \lambda(f(\widetilde{x}) - f(a)) \leqslant \frac{2M}{r}\|x-a\|_1.$$

另一方面, 由 λ 的定义可知 (图 2.17-1)

$$a = \frac{1}{1+\lambda}x + \frac{\lambda}{1+\lambda}\hat{x},$$

其中 $\hat{x} := a - (\tilde{x} - a)$. 因此, $f(a) \leqslant \frac{1}{1+\lambda}f(x) + \frac{\lambda}{1+\lambda}f(\hat{x})$. 由此可知, 对所有的 $x \in B$,

$$f(a) - f(x) \leqslant \lambda(f(\hat{x}) - f(a)) \leqslant \frac{2M}{r}\|x-a\|_1.$$ □

注 还可以证明 f 是局部 Lipschitz 连续的, 见习题 2.17-11.

值得指出的是定理 2.17-1 在无限维空间中并不成立. 例如, 考察空间 $\mathcal{P}$, 赋以范数 $\|p\| = \sup_{0\leqslant x\leqslant 1}|p(x)|$, 线性泛函 $\ell : p \in \mathcal{P} \to p(3)$, 这是一个凸函数. 但是 ℓ 是不连续的: 考察序列 $(p_n)_{n=1}^{\infty}$, 其中 $p_n(x) = \left(\frac{x}{2}\right)^n, x \in \mathbb{R}$, 则有 $\|p_n\| \xrightarrow[n\to\infty]{} 0$, 但 $\ell(p_n) \xrightarrow[n\to\infty]{} \infty$.

凸函数的进一步的例子在习题 2.17-1 和 2.17-2 中给出.

赋范向量空间 $(X, \|\cdot\|)$ 称为*严格凸的*或 rotund, 是指下述较强的性质成立:

$$\|x\| = \|y\| = 1 \text{ 且当 } x \neq y \text{ 时}, \left\|\frac{x+y}{2}\right\| < 1.$$

赋范向量空间 $(X, \|\cdot\|)$ 称为一致凸的, 是指下述更强的性质成立: 对任何 $\varepsilon > 0$, 存在 $\delta = \delta(\varepsilon) > 0$, 使得

$$\|x\| = \|y\| = 1, \text{ 且当 } \|x - y\| \geqslant \varepsilon \text{ 时}, \left\|\frac{x+y}{2}\right\| \leqslant 1 - \delta(\varepsilon).$$

下面将会看到一致凸的概念在对弱收敛 (定理 5.12-3) 和自反空间 (定理 5.14-3) 的分析时特别重要.

对 $1 < p < \infty$, 空间 ℓ^p 和 $L^p(\Omega)$ (2.4 和 2.5 节) 即提供了一致凸空间的基本例子 (习题 2.17-8 和 2.17-9).

习题

2.17-1　设 $(X, \|\cdot\|)$ 为赋范向量空间. 证明对任何 $p \geqslant 1$, 函数 $f: x \in X \to \|x\|^p$ 是凸的.

2.17-2　设 X 为实向量空间, A 为 X 中的凸子集.

(1) 证明函数 $f: A \to \mathbb{R}$ 是凸的, 当且仅当其定义为

$$E_{\rm pi} f := \{(x, y) \in X \times \mathbb{R}; x \in A, y \geqslant f(x)\}$$

的上境图 (epigraph) 是向量空间 $X \times \mathbb{R}$ 中的凸子集.

注　此性质对在 $\mathbb{R} \cup \{\infty\}$ 上取值的凸函数亦成立, 见定理 9.2-1.

(2) 设有一族凸函数 $(f_i)_{i\in I}$, 其中凸函数 $f_i: A \to \mathbb{R}$ 满足对一切 $x \in A, \sup_{i\in I} f_i(x) < \infty$. 证明函数 $\sup_{i\in I} f_i: A \to \mathbb{R}$ 也是凸的.

2.17-3　设 $(X, \|\cdot\|)$ 为赋范向量空间, $f: X \to \mathbb{R}$ 为凸函数, 且有一个局部极小值点 $x_0 \in X$, 即存在 $r > 0$, 使得当 $x \in B(x_0, r)$ 时, $f(x) \geqslant f(x_0)$. 证明 x_0 是 f 的整体最小值点, 即对任意的 $x \in X$, 均有 $f(x) \geqslant f(x_0)$.

2.17-4　设 X 为向量空间, $f: X \to \mathbb{R}$ 为严格凸函数, 证明 f 至多只有一个最小值点; 如果 f 有一个最小值点 x_0, 则它一定是严格最小的, 即对一切 $x \in X$, $x \neq x_0$, 均有 $f(x) > f(x_0)$.

2.17-5　设 X 为有限维赋范向量空间, $f: X \to \mathbb{R}$ 为凸函数, 且有严格整体最小值点 $x_0 \in X$, 即对一切 $x \in X$, $x \neq x_0$, $f(x) > f(x_0)$, 证明当 $\|x\| \to \infty$ 时一致地有 $f(x) \to \infty$, 即

$$\lim_{r\to\infty} \big(\inf_{\|x\|\geqslant r} f(x) \big) = \infty.$$

2.17-6　证明有限维严格凸空间是一致凸的.

2.17-7　证明空间 ℓ^1 和 ℓ^∞, 以及空间 $L^1(\Omega)$ 和 $L^\infty(\Omega)$ 均非严格凸的.

2.17-8　以下给定数 $1 < p \leqslant 2$, $q > 1$ 定义为 $\frac{1}{p} + \frac{1}{q} = 1$.

(1) 证明 $(1+t)^q + (1-t)^q \leqslant 2(1+t^p)^{q/p}$ 对一切 $0 \leqslant t \leqslant 1$ 成立.

(2) 利用 (1) 证明对一切 $\alpha, \beta \in \mathbb{K}, |\alpha+\beta|^q + |\alpha-\beta|^q \leqslant 2(|\alpha|^p + |\beta|^p)^{q/p}$.

(3) 利用 (2) 证明下述的 Clarkson 不等式[38). 当 $1 < p \leqslant 2$ 时, 对一切 $x, y \in \ell^p$, 有

$$(\|x+y\|_{\ell^p})^q + (\|x-y\|_{\ell^p})^q \leqslant 2((\|x\|_{\ell^p})^p + (\|y\|_{\ell^p})^p)^{q/p};$$

对一切 $f, g \in L^p(\Omega)$,

$$(\|f+g\|_{L^p(\Omega)})^q + (\|f-g\|_{L^p(\Omega)})^q \leqslant 2((\|f\|_{L^p(\Omega)})^p + (\|g\|_{L^p(\Omega)})^p)^{\frac{q}{p}}.$$

(4) 证明当 $p = 2$(此时 $q = 2$) 时, Clarkson 不等式成为等式; 这个等式是在任何内积空间中的成立的平行四边形公式 (定理 4.1-2) 的特例.

(5) 综上证明: 当 $1 < p \leqslant 2$ 时空间 ℓ^p 和 $L^p(\Omega)$ 均为一致凸的.

2.17-9 以下给定数 $p \geqslant 2$.

(1) 证明当 $0 \leqslant t \leqslant 1$ 时

$$\left(\frac{1+t}{2}\right)^p + \left(\frac{1-t}{2}\right)^p \leqslant \frac{1}{2}(1+t^p).$$

(2) 利用 (1) 证明对任何 $\alpha, \beta \in \mathbb{K}$, $\left|\frac{\alpha+\beta}{2}\right|^p + \left|\frac{\alpha-\beta}{2}\right|^p \leqslant \frac{|\alpha|^p}{2} + \frac{|\beta|^p}{2}$.

(3) 利用 (2) 证明下述的 Clarkson 不等式对 $p \geqslant 2$ 成立 (在习题 2.17-8(4) 中的讨论也适用于这个不等式): 对一切 $x, y \in \ell^p$, 有

$$(\|x+y\|_{\ell^p})^p + (\|x-y\|_{\ell^p})^p \leqslant 2^{p-1}((\|x\|_{\ell^p})^p + (\|y\|_{\ell^p})^p);$$

对一切 $f, g \in L^p(\Omega)$,

$$(\|f+g\|_{L^p(\Omega)})^p + (\|f-g\|_{L^p(\Omega)})^p \leqslant 2^{p-1}((\|f\|_{L^p(\Omega)})^p + (\|g\|_{L^p(\Omega)})^p).$$

(4) 综上证明: 对 $p \geqslant 2$, 空间 ℓ^p 和 $L^p(\Omega)$ 均为一致凸的.

2.17-10 设 $\varphi : \mathbb{R} \to \mathbb{R}$ 为凸函数 (因而由定理 2.17-1, 它是连续的).

(1) 证明对任何整数 $m \geqslant 1$ 和任何 $\zeta_i \in \mathbb{R}, \lambda_i > 0,\ 1 \leqslant i \leqslant m$,

$$\varphi\left(\frac{\sum_{i=1}^m \lambda_i \zeta_i}{\sum_{i=1}^m \lambda_i}\right) \leqslant \frac{\sum_{i=1}^m \lambda_i \varphi(\zeta_i)}{\sum_{i=1}^m \lambda_i}.$$

(2) 利用 (1) 和函数 $x \in]0, \infty[\to -\log x$ 的凸性, 证明算术几何平均不等式, 即对任何 $\zeta_i > 0,\ 1 \leqslant i \leqslant m$,

$$\left(\prod_{i=1}^m \zeta_i\right)^{\frac{1}{m}} \leqslant \frac{1}{m}\sum_{i=1}^m \zeta_i.$$

(3) 证明对 $\mathbb{R}^n$ 的任何开子集 Ω 和任何非负函数 $f \in L^1(\Omega)$,

$$\varphi\left(\frac{1}{\operatorname{meas}\Omega}\int_\Omega f(x)dx\right) \leqslant \frac{1}{\operatorname{meas}\Omega}\int_\Omega \varphi(f(x))dx.$$

[38) J. A. Clarkson [1936]: Uniformly convex spaces. Transactions of the American Mathematical Society 40, 396–414.

(1) 和 (2) 中的不等式称之为 Jensen 不等式[39); 在空间 ℓ^p 中也有 Jensen 不等式 (见习题 2.4-4).

2.17-11 假设与记号均同于定理 2.17-1. 证明: 对任意给定的点 $a \in \Omega$, 存在邻域 $V(a)$ 和常数 $C = C(a, V(a)) > 0$, 使得当 $x, y \in V(a)$ 时, $|f(x) - f(y)| \leqslant C\|x - y\|_1$.

[39) J. L. W. V. Jensen [1906]: Sur les fonctions convexes et les inégalités entre les valeurs moyennes. Acta Mathematica 30, 175–193.

第 3 章 Banach 空间

引言

Banach 空间, 即完备的赋范向量空间在线性与非线性泛函分析中起着核心的作用. 本章旨在建立起它们最重要的基本性质.

作为开始, 我们将给出并研究在本书以下各章节将经常出现的 Banach 空间的基本例子, 如赋以 sup 范数的由紧集 K 到 Banach 空间 Y 的所有连续函数组成的空间 $\mathcal{C}(K;Y)$ (3.2 节), 空间 ℓ^p 和 $L^p(\Omega)$, $1 \leqslant p \leqslant \infty$ (3.4 节), Y 为 Banach 空间时的空间 $\mathcal{L}(X;Y)$, 作为特殊情况下众多重要的对偶空间 (3.5 节). 特别地, 给出了空间 $L^p(\Omega)$, $1 \leqslant p < \infty$ 的基本的 F. Riesz *表示定理* (定理 3.5-3 的证明), 这类空间恒同于其对偶空间.

在赋范向量空间完备的假设下可以证明一系列意义深远的结果, 本章将建立并应用其中三个结果 (意义更深远但性质更精致的结果将在第 5 章中建立).

第一个结果是在 Banach 空间中定义收敛级数成为可能 (3.6 节). 例如, 当线性算子 A 作用于 Banach 空间, 且 $\|A\| < 1$ 时, 利用 Neumann *级数*可以计算形为 $I - A$ 的算子的逆 (定理 3.6-2).

第二个结果, 也许是 Banach 空间理论最基本的结果, 是 Banach *不动点定理* (定理 3.7-1). 其重要性在本章中显示于两个应用, 第一是对非线性常微分方程, 如单摆方程的 Cauchy-Lipschitz *定理* (定理 3.8-1), 其二是对*非线性的两点边值问题* (定理 3.9-1). 后面还将用到 Banach 不动点定理, 如作为基本的*隐函数定理* (7 章) 的基石, 这些应用显示 Banach 不动点定理也是*非线性泛函分析*的一个基本定理 (实际上在本书中第一个出现的这类结果), 也许是最基本的一个定理.

第三个结果也是一个基本定理, 即 Ascoli-Arzelà *定理*, 它给出了当 K 为紧集

时, 空间 $\mathcal{C}(K;\mathbb{R})$ 中紧集的特征 (定理 3.10-1). 其重要性可见于非线性常微分方程的 Cauchy-Peano 定理 (定理 3.11-1).

3.1 Banach 空间; 基本性质

赋范向量空间 $(X,\|\cdot\|)$ 称为 Banach 空间[1], 是指距离空间 (X,d) 是完备的, 这里 X 上的距离 d 定义为 $d(x,y):=\|x-y\|$ (定理 2.2-1).

Banach 空间因袭了完备距离空间的所有性质, 如那些在第 1.12 节所列的性质. 此外, 由于它还有赋范向量空间这一更丰富的结构, 所以上述若干性质还可进一步深化. 例如, 在一个稠密子集上定义且连续的一致连续映射, 如果取值于一个完备空间, 则可唯一地连续延拓到全空间 (定理 1.12-3), 将这一结果应用于赋范向量空间上的线性映射, 在定理 3.1-1 中就给出了一个特殊的形式. 考虑到这个结果的重要性, 我们在这里给出一个独立的证明 (即证明的第一部分不依赖定理 1.12-3).

定理 3.1-1 (唯一连续线性延拓) 设 X 为赋范向量空间 $\widetilde{X}$ 的稠密子空间, Y 为 Banach 空间, $A:X\to Y$ 为连续线性算子, 则存在唯一的连续线性算子 $\widetilde{A}:\widetilde{X}\to Y$, $\widetilde{A}$ 是 A 的延拓, 即当 $x\in X$ 时 $\widetilde{A}x=Ax$; 而且, 对任何 $\widetilde{x}\in\widetilde{X}$ 和元素 $x_n\in X$ 的任何序列 $(x_n)_{n=1}^{\infty}$, 当它在 $\widetilde{X}$ 中满足 $\lim_{n\to\infty}x_n=\widetilde{x}$ 时

$$\widetilde{A}\widetilde{x}:=\lim_{n\to\infty}Ax_n.$$

另外,

$$\|\widetilde{A}\|_{\mathcal{L}(\widetilde{X};Y)}=\|A\|_{\mathcal{L}(X;Y)}.$$

证明 (i) 首先要定义这样的延拓.

对任意给定的 $\widetilde{x}\in\widetilde{X}=\overline{X}$, 取向量 $x_n\in X$ 的序列 $(x_n)_{n=1}^{\infty}$, 使之收敛于 $\widetilde{x}$. 因为对一切 $m,n\geqslant 1$,

$$\|Ax_m-Ax_n\|\leqslant\|A\|\,\|x_m-x_n\|,$$

而 Y 是完备的, 因此存在 $y\in Y$, 使得当 $n\to\infty$ 时 $Ax_n\to y$ (这里实际上用到 A 的一致连续性, 见定理 2.9-3(a)). 另外, y 不依赖收敛于 $\widetilde{x}$ 的向量 $x_n\in X$ 的特殊序列的选取. 为此, 考察另一个这样的序列 $(x_n')_{n=1}^{\infty}$, 这时, 序列 $(Ax_n)_{n=1}^{\infty}$ 和 $(Ax_n')_{n=1}^{\infty}$ 作为同一个 Cauchy 序列 $(Ax_1,Ax_1',Ax_2,Ax_2',\cdots)$ 的子列必有同一个极限.

[1] Banach 空间的命名是为纪念波兰数学家 Stefan Banach (1892—1945). 他本质上创建了这一理论, 并且在一本数学史上最有影响的著作之一 Banach [1932] 中详尽地作了展开 (关于传记和历史性的叙述可见 Pietsch [2007] 和 Jakimowicz & Miranovicz [2011]). 他和其他一些数学家, 如 Karol Borsuk (1905—1982), Stanislaw Saks (1897—1942), Juliusz Schauder (1899—1943) 及 Hugo Steinhaus (1887—1972) (所有这些名字将在本书中后面出现) 经常在 Lwów (当时的波兰东部, 现在的乌克兰西部的 Lviv) 的 “苏格兰咖啡馆” 里工作, 后者是传说中过去的那个数学时代的象征, 见 Maudlin [1981].

令 $\widetilde{A}\widetilde{x} := y$, 这就定义了一个映射 $\widetilde{A}: \widetilde{X} \to Y$, 它显然是 A 的延拓 (如果 $x \in X$, 则可用特殊的序列 $(x, x, \cdots)$ 定义 $\widetilde{A}x$).

(ii) 由 (i) 定义的 $A \in \mathcal{L}(X;Y)$ 的延拓 $\widetilde{A}: \widetilde{X} \to Y$ 是连续映射, $\widetilde{A}$ 是 A 到 $\widetilde{X}$ 的唯一连续延拓.

不失一般性, 设 $A \neq 0$. 给定 $\varepsilon > 0$, 存在 $\delta := \frac{\varepsilon}{2\|A\|}$, 当 $x, x' \in X, \|x - x'\| \leqslant 2\delta$ 时, $\|Ax - Ax'\| \leqslant \varepsilon$. 设 $\widetilde{x}, \widetilde{x}' \in \widetilde{X}$ 满足 $\|\widetilde{x} - \widetilde{x}'\| \leqslant \delta$, 取 $x_n, x_n' \in X$, $n \geqslant 1$, 使得当 $n \to \infty$ 时 $x_n \to \widetilde{x}, x_n' \to \widetilde{x}'$, 于是, 存在 $n_0 = n_0(\widetilde{x}, \widetilde{x}')$, 使当 $n \geqslant n_0$ 时 $\|x_n - x_n'\| \leqslant 2\delta$. 于是, 当 $n \geqslant n_0$ 时 $\|Ax_n - Ax_n'\| \leqslant \varepsilon$. 这样

$$\|\widetilde{A}\widetilde{x} - \widetilde{A}\widetilde{x}'\| = \lim_{n\to\infty} \|Ax_n - Ax_n'\| \leqslant \varepsilon$$

对任何满足 $\|\widetilde{x} - \widetilde{x}'\| \leqslant \delta$ 的 $\widetilde{x}, \widetilde{x}' \in \widetilde{X}$ 成立. 这就证明了 A 的延拓 $\widetilde{A}: \widetilde{X} \to Y$ 是连续的 (实际上还是一致连续的).

设 $\widetilde{A}': \widetilde{X} \to Y$ 是 A 的另一个连续延拓. 给定 $\widetilde{x} \in \widetilde{X} - X$, 令 $x_n \in X, n \geqslant 1$, 满足当 $n \to \infty$ 时 $x_n \to \widetilde{x}$. 于是, 由 $\widetilde{A}x$ 的定义和 $\widetilde{A}'$ 的连续性, 可得

$$\widetilde{A}x = \lim_{n\to\infty} Ax_n, \quad \widetilde{A}'x = \lim_{n\to\infty} \widetilde{A}'x_n = \lim_{n\to\infty} Ax_n.$$

由赋范向量空间上序列极限的唯一性 (定理 1.10-1), 即得 $\widetilde{A}\widetilde{x} = \widetilde{A}'\widetilde{x}$. 因此, $\widetilde{A}$ 是 A 的到 $\widetilde{X}$ 上的唯一连续延拓.

(iii) 连续映射 $\widetilde{A}$ 也是线性算子, 而且 $\|\widetilde{A}\|_{\mathcal{L}(\widetilde{X};Y)} = \|A\|_{\mathcal{L}(X;Y)}$.

对任意给定的 $\widetilde{x}, \widetilde{x}' \in \widetilde{X}$, 取 $x_n \in X, x_n' \in X, n \geqslant 1$, 使得当 $n \to \infty$ 时, $x_n \to \widetilde{x}, x_n' \to \widetilde{x}'$. 于是, 对任意给定的 $\alpha, \alpha' \in \mathbb{K}$, 由赋范向量空间中加法与数乘的连续性 (定理 2.2-5), 可得

$$\begin{aligned}\widetilde{A}(\alpha\widetilde{x} + \alpha'\widetilde{x}') &= \lim_{n\to\infty} (A(\alpha x_n + \alpha' x_n')) \\ &= \lim_{n\to\infty} (\alpha A x_n + \alpha' A x_n') = \alpha\widetilde{A}\widetilde{x} + \alpha'\widetilde{A}\widetilde{x}'.\end{aligned}$$

因 $X \subset \widetilde{X}$, 显然有 $\|A\|_{\mathcal{L}(X,Y)} \leqslant \|\widetilde{A}\|_{\mathcal{L}(\widetilde{X};Y)}$. 对任意给定的 $\widetilde{x} \in \widetilde{X}$, 取 $x_n \in X, n \geqslant 1$, 使得当 $n \to \infty$ 时 $x_n \to \widetilde{x}$. 于是, 当 $n \geqslant 1$ 时

$$\|Ax_n\| \leqslant \|A\|\,\|x_n\|, \quad \|\widetilde{A}\widetilde{x}\| = \lim_{n\to\infty} \|Ax_n\|.$$

于是, 由当 $n \to \infty$ 时 $\|x_n\| \to \|\widetilde{x}\|$, 得到 $\|\widetilde{A}\widetilde{x}\| \leqslant \|A\|\,\|\widetilde{x}\|$. 因此, $\|\widetilde{A}\|_{\mathcal{L}(\widetilde{X},Y)} \leqslant \|A\|_{\mathcal{L}(X;Y)}$. □

这类唯一连续线性延拓的一个重要的例子是 "迹算子", 由它可以对 Sobolev 空间中的函数定义 "边界值" (6.6 节).

距离空间完备化的过程 (定理 1.12-4) 提供了结果可以在赋范向量空间中进一步深化的另一个例子. 同样是出于其重要性的考虑, 我们也对这个结果给出一个独立的证明. 作为一个有趣的补充, 可参见习题 3.1-1.

首先, 我们需要一个定义. 设线性算子 σ 是赋范向量空间 $(X;\|\cdot\|_X)$ 到赋范向量空间 $(Y,\|\cdot\|_Y)$ 的, 或相应地, 到 $(Y,\|\cdot\|_Y)$ 上的映射, 满足

$$\|\sigma x\|_Y = \|x\|_X$$

对所有 $x \in X$ 成立, 则称 σ 为 X 到 Y 的线性等距算子, 或相应地, X 到 Y 上的线性等距算子. 线性等距算子显然是单射, 而且是连续的.

定理 3.1-2 (赋范向量空间的完备化) 设 $(X,\|\cdot\|_X)$ 是 $\mathbb{K}$ 上的赋范向量空间, 则存在 $\mathbb{K}$ 上的 Banach 空间 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 和线性等距算子 $\sigma : X \to \widetilde{X}$, 使得 $\sigma(X)$ 在 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 中稠密.

另外, 如果 $(\hat{X},\|\cdot\|_{\widehat{X}})$ 是 $\mathbb{K}$ 上的 Banach 空间, 而且存在 X 到 $\widehat{X}$ 的一个稠密子集上的线性等距算子, 则必存在从 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 到 $(\widehat{X},\|\cdot\|_{\widehat{X}})$ 上的线性等距算子.

这样, 称为 $(X,\|\cdot\|_X)$ 的完备化空间 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 在线性等距双射的意义下是唯一的. 作为赋范向量空间, 空间 X 可与其完备化空间 $\widetilde{X}$ 中的一个稠密子集视为同一.

证明 (i) 构造完备化空间.

为记号便利起见, 以下记序列 $(x_n)_{n=1}^{\infty}$ 为 (x_n). 在由向量 $x_n \in X$ 的所有 Cauchy 序列组成的集合 $\mathcal{C}$ 上, 容易验证由 $(x_n) \sim (y_n)$ 当且仅当 $n \to \infty$ 时 $\|x_n - y_n\| \to 0$ 定义了一个等价关系 $\mathcal{R}$.

首先, 我们在商集 $\widetilde{X} := \mathcal{C}/\mathcal{R}$ 上可以自然地引入加法和数乘并赋以一个范数, 使之成为同一个域 $\mathbb{K}$ 上的赋范向量空间. 用 $[(x_n)]$ 表示 (x_n) 的等价类 (等价关系、等价类和商集的概念可见 1.1 节). 令

$$\begin{aligned}
0 &:= [(x_n)], \text{ 这里对一切 } n \geqslant 1,\ x_n = 0,\\
[(x_n)] + [(y_n)] &:= [(x_n + y_n)],\\
\alpha[(x_n)] &:= [(\alpha x_n)], \quad \alpha \in \mathbb{K},\\
\|[(x_n)]\|_{\widetilde{X}} &:= \lim_{n\to\infty} \|x_n\|.
\end{aligned}$$

为验证这里加法和数乘的定义有意义, 即需说明它们与等价类中特殊的 Cauchy 序列的选取无关, 注意到如果 $(x_n) \sim (x'_n), (y_n) \sim (y'_n)$, 则 $(x_n + y_n)$ 与 $(x'_n + y'_n)$ 也是 Cauchy 序列, 且因

$$\|(x_n + y_n) - (x'_n + y'_n)\| \leqslant \|x_n - x'_n\| + \|y_n - y'_n\|,$$

故 $(x_n + y_n) \sim (x'_n + y'_n)$. 类似地, 如果 $(x_n) \sim (x'_n)$, 则 (αx_n) 和 $(\alpha x'_n)$ 也是 Cauchy 序列, 且因 $\|\alpha x_n - \alpha x'_n\| = |\alpha|\,\|x_n - x'_n\|$, 故而 $(\alpha x_n) \sim (\alpha x'_n)$.

由不等式 $|\,\|x_n\| - \|x_m\|\,| \leqslant \|x_n - x_m\| \leqslant \|x_n - x_m\|$ 可知, 如果 (x_n) 是 X 中向量的 Cauchy 序列, 则 $(\|x_n\|)$ 是实数的 Cauchy 序列. 因为 $(\mathbb{R},|\cdot|)$ 是完备的, 所以 $\lim_{n\to\infty}\|x_n\|$ 是一个确定的实数. 另外, 如果 $(x_n) \sim (x'_n)$, 则由不等

式 $|\|x_n\|-\|x_n'\||\leqslant\|x_n-x_n'\|$ 可知 $\lim_{n\to\infty}\|x_n\|=\lim_{n\to\infty}\|x_n'\|$. 因此, 数 $\|[(x_n)]\|_{\widetilde{X}}$ 与 $[(x_n)]$ 中特殊 Cauchy 序列的选取无关.

类似地, 可以验证映射 $\|\cdot\|_{\widetilde{X}}:\widetilde{X}\to\mathbb{R}$ 是 $\widetilde{X}$ 上的范数.

(ii) 对任意给定的 $x\in X$, 用 $\sigma(x)\in\widetilde{X}$ 表示对应于特殊的 Cauchy 序列 (x_n) 的等价类, 这里当 $n\geqslant 1$ 时 $x_n=x$. 可以直接验证这样定义的映射 $\sigma:X\to\widetilde{X}$ 是 X 到 $\widetilde{X}$ 的线性等距算子.

现在我们来证明直接像 $\sigma(X)$ 在 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 中稠密. 任意给定 $\widetilde{x}=[(x_n)]\in\widetilde{X}$ 和任意的 $\varepsilon>0$. 因为 (x_n) 是 Cauchy 序列, 所以存在整数 $n_0=n_0(\widetilde{x},\varepsilon)\geqslant 1$, 使得当 $n\geqslant n_0$ 时, $\|x_n-x_{n_0}\|\leqslant\varepsilon$. 令 $\widetilde{x}_0:=[(y_n)]$ 这里对所有的 $n\geqslant 1, y_n=x_{n_0}$. 因为 $\widetilde{x}_0=\sigma(x_{n_0})$, 故 $\widetilde{x}_0\in\sigma(X)$, 而且

$$\|\widetilde{x}-\widetilde{x}_0\|_{\widetilde{X}}=\lim_{n\to\infty}\|\widetilde{x}_n-\widetilde{x}_{n_0}\|_{\widetilde{X}}\leqslant\varepsilon.$$

因此, $\sigma(X)$ 在 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 中稠密.

(iii) 赋范向量空间 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 是完备的. 设 $(\widetilde{x}^k)_{k=1}^{\infty}$ 是 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 中的 Cauchy 序列, 由 (ii), $\sigma(X)$ 在 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 中稠密, 所以对每个 $k\geqslant 1$, 存在 $x^k\in X$, 使得 $\|\widetilde{x}^k-\sigma(x^k)\|_{\widetilde{x}}\leqslant\frac{1}{k}$, $(x^k)_{k=1}^{\infty}$ 是 X 中的 Cauchy 序列, 这是因为由 (ii), $\sigma:X\to\widetilde{X}$ 是等距算子, 故对所有的 $k,\ell\geqslant 1$,

$$\|x^k-x^\ell\|_X=\|\sigma(x^k)-\sigma(x^\ell)\|_{\widetilde{X}}.$$

因此,

$$\begin{aligned}\|x^k-x^\ell\|_X&=\|\widetilde{x}^k-\sigma(x^k)\|_{\widetilde{X}}+\|\widetilde{x}^\ell-\sigma(x^\ell)\|_{\widetilde{X}}+\|\widetilde{x}^k-\widetilde{x}^\ell\|_{\widetilde{X}}\\&\leqslant\frac{1}{k}+\frac{1}{\ell}+\|\widetilde{x}^k-\widetilde{x}^\ell\|_{\widetilde{X}}.\end{aligned}$$

令 $\widetilde{x}:=[(x^k)]$, 则当 $k\to\infty$ 时 $\|\widetilde{x}^k-\widetilde{x}\|_{\widetilde{X}}\to 0$, 为证明这一点, 注意到

$$\begin{aligned}\|\widetilde{x}^k-\widetilde{x}\|_{\widetilde{X}}&\leqslant\|\widetilde{x}^k-\sigma(x^k)\|_{\widetilde{X}}+\|\sigma(x^k)-\widetilde{x}\|_{\widetilde{X}}\\&\leqslant\frac{1}{k}+\|\sigma(x^k)-\widetilde{x}\|_{\widetilde{X}},\end{aligned}$$

因为 $\sigma(x^k)=[(x^k,x^k,\cdots,x^k,\cdots)]$, $\widetilde{x}=[(x^1,x^2,\cdots,x^k,x^{k+1},\cdots)]$, 由 $\|\cdot\|_{\widetilde{X}}$ 的定义

$$\|\sigma(x^k)-\widetilde{x}\|_{\widetilde{X}}=\lim_{n\to\infty}\|x^k-x^n\|=0.$$

因此,

$$\lim_{k\to\infty}\|\sigma(x^k)-\widetilde{x}\|_{\widetilde{X}}=0,$$

即得 $\lim_{k\to\infty}\|\widetilde{x}^k-\widetilde{x}\|_{\widetilde{X}}=0$, 这就证明了 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 也是完备的.

(iv) 再设存在线性等距算子 $\tau:X\to\widehat{X}$, 这里 $(\widehat{X},\|\cdot\|_{\widehat{X}})$ 为 Banach 空间, $\tau(X)$ 是 $\widehat{X}$ 的稠密子集.

于是 $\tau\circ\sigma^{-1}:\sigma(X)\to\widehat{X}$ 是 $(\sigma(X),\|\cdot\|_{\widetilde{X}})$ 到 $(\widehat{X},\|\cdot\|_{\widehat{X}})$ 的线性等距算子. 因为 $\sigma(X)$ 在 $\widetilde{X}$ 中稠密, $\widehat{X}$ 是完备的. 由连续线性延拓的唯一性定理 (定理 3.1-1), $\tau\circ\sigma^{-1}$ 存在唯一的连续线性延拓 $\varphi:\widetilde{X}\to\widehat{X}$, 显然, φ 也是线性等距算子 (为此, 可考察 $\sigma(X)$ 中的序列, 并利用范数的连续性).

类似地, 线性等距算子 $\sigma\circ\tau^{-1}:\tau(X)\to\widetilde{X}$ 也有唯一连续线性延拓 $\psi:\widehat{X}\to\widetilde{X}$, 它是 $\widehat{X}$ 到 $\widetilde{X}$ 的线性等距算子.

由构造, 线性等距算子 $\varphi\circ\psi:\widehat{X}\to\widehat{X}$ 在 $\tau(X)$ 上的限制即恒等映射 $I_{\tau(X)}$. 再利用连续线性延拓唯一性的定理可知 $\varphi\circ\psi=I_{\widehat{X}}$.

因此, 线性等距算子 $\varphi:\widetilde{X}\to\widehat{X}$ 不仅是单射, 而且是满射 (因为对任何 $\widehat{x}\in\widehat{X}$, $\varphi(\psi(\widehat{x}))=\widehat{x}$). 从而 φ 是 $(\widetilde{X},\|\cdot\|_{\widetilde{X}})$ 到 $(\widehat{X},\|\cdot\|_{\widehat{X}})$ 上的线性等距算子, 证毕. □

设 Ω 是 $\mathbb{R}^n$ 中的开子集, $1\leqslant p<\infty$, Lebesgue 空间 $L^p(\Omega)$ 是完备的 (将在 3.4 节中证明), 它们提供了基本的完备化空间的例子. 例如, 当 Ω 有界时 (它保证了当 $f\in\mathcal{C}(\overline{\Omega})$ 时 $\|f\|_{L^p(\Omega)}<\infty$), 作为空间 $(\mathcal{C}(\overline{\Omega}),\|\cdot\|_{L^p(\Omega)})$ 的完备化空间; 或一般地, 作为空间 $(\mathcal{C}_{\mathcal{C}}(\Omega),\|\cdot\|_{L^p(\Omega)})$ 的完备化空间, 这里 $\mathcal{C}_{\mathcal{C}}(\Omega)$ 是所有在 Ω 上连续, 且在 Ω 中有紧支集的函数全体组成的空间; 或者一般地, 作为空间 $(\mathcal{D}(\Omega),\|\cdot\|_{L^p(\Omega)})$ 的完备化空间, 这里的 $\mathcal{D}(\Omega)$ 表示所有在 Ω 上无限阶可数, 且在 Ω 中有紧支集的函数全体组成的空间 (定理 2.5-3 和 2.6-2) 显然, 最后一个空间的稠密性可以导出另外两种情况的结果.

注 定理 3.1-2 证明中完备空间的构造借鉴了从集合 $\mathbb{Q}$ 到完备赋范向量空间 $(\mathbb{R},|\cdot|)$ 的构造, 即利用有理数的 Cauchy 序列的等价类的方法. 然而, 也应注意到上述证明中本质上用到了赋范向量空间 $(\mathbb{R},|\cdot|)$ 的完备性 (见第 (i) 部分).

我们以 Banach 空间的三个一般性质来结束本节. 前面几个性质虽然几乎是显然的, 但仍值得一提.

定理 3.1-3 设 X 为 Banach 空间, Y 为赋范向量空间, $A\in\mathcal{L}(X;Y)$ 为双射, 满足 $A^{-1}\in\mathcal{L}(Y;X)$. 则 Y 是 Banach 空间.

证明 设 (y_n) 是 Y 中的 Cauchy 序列, 则 $(A^{-1}y_n)$ 是 X 中的 Cauchy 序列 (因为 $A^{-1}\in\mathcal{L}(Y;X)$), 它收敛于极限 $x\in X$(因为 X 是完备的). 因此, $y_n=A(A^{-1}y_n)$ 在 Y 中收敛于 $Ax\in Y$(因为 $A\in\mathcal{L}(X;Y)$). □

定理 3.1-4 设 X 为 Banach 空间, Y 为赋范向量空间, $A\in\mathcal{L}(X;Y)$. 如果存在常数 C, 使得对所有的 $x\in X$,

$$\|x\|\leqslant C\|Ax\|,$$

则 $\operatorname{Im}A$ 也是 Banach 空间, 从而是 Y 的闭子空间.

证明 设 $y_n=Ax_n\in\operatorname{Im}A$, $n\geqslant1$, 使 (y_n) 是 Y 中的 Cauchy 序列, 由条件中的不等式可知 (x_n) 是 X 中的 Cauchy 序列, 因为 X 是完备的, 从而收敛于极限 $x\in X$.

由 A 的连续性, 当 $n \to \infty$ 时

$$y_n = Ax_n \to y := Ax$$

因此 $y = Ax \in \operatorname{Im} A$. 说明 Y 的子空间 $\operatorname{Im} A$ 是完备的. 特别地, $\operatorname{Im} A$ 在 Y 中是闭的 (定理 1.12-2(a)). □

注 定理 3.1-4 条件中的不等式说明 A 是单射, 而且 $A : X \to \operatorname{Im} A$ 的逆映射是 $\operatorname{Im} A$ 到 X 上的连续线性算子 (定理 2.9-4).

第三个结果 (与前两条不同, 其证明颇为不易) 是 Banach 空间一个十分重要的性质. 例如, 它在 Schauder 不动点定理 (定理 9.12-1) 的证明中起着关键的作用. 在 2.16 节中已经定义过闭凸包.

定理 3.1-5 *Banach 空间的紧子集的闭凸包也是紧的.*

证明 设 A 为 Banach 空间 X 的紧子集, 则作为完备距离空间的闭子集, $\overline{\text{co}}\, A$ 也是完备距离空间, 这就是 X 为 Banach 空间的假设是必要的原因. 于是, 由定理 1.13-3, 只要证明 $\overline{\text{co}}\, A = \overline{\text{co}\, A}$ (定理 2.16-3) 是准紧的, 或等价地, 证明 $\text{co}\, A$ 是准紧的即可.

对任意给定的 $\varepsilon > 0$, 由 A 为紧集的假设, 存在 A 的有限子集 $A(\varepsilon)$, 使得

$$A \subset \bigcup_{x \in A(\varepsilon)} B\left(x; \frac{\varepsilon}{2}\right).$$

因为 $A(\varepsilon)$ 为有限集, 其凸包 $\text{co}\, A(\varepsilon)$ 为紧集 (定理 2.16-2), 从而是准紧的, 所以存在数 $m = m(\varepsilon)$ 和有限个点 $y_i = y_i(\varepsilon) \in \text{co}\, A(\varepsilon), 1 \leqslant i \leqslant m$, 使得

$$\text{co}\, A(\varepsilon) \subset \bigcup_{i=1}^{m} B\left(y_i, \frac{\varepsilon}{2}\right).$$

任意给定点 $y \in \text{co}\, A$, 存在足标的有限集 $J = J(y)$, 使得 (定理 2.16-1).

$$y = \sum_{j \in J} \lambda_j x_j, \text{ 其中 } x_j \in A, \lambda_j \geqslant 0, j \in J, \sum_{j \in J} \lambda_j = 1.$$

对每个 $j \in J$, 存在点 $x(j) \in A(\varepsilon)$, 使得 $x_j \in B\left(x(j); \frac{\varepsilon}{2}\right)$, 即 $\|x_j - x(j)\| < \frac{\varepsilon}{2}$. 于是, 点

$$z := \sum_{j \in J} \lambda_j x(j)$$

满足

$$z \in \text{co}\, A(\varepsilon), \text{ 且 } \|y - z\| = \left\| \sum_{j \in J} \lambda_j (x_j - x(j)) \right\| < \frac{\varepsilon}{2}.$$

因为 $z \in \text{co}\, A(\varepsilon)$, 故而存在整数 $i = i(z)$, 其中 $1 \leqslant i \leqslant m$, 使得 $z \in B\left(y_i; \frac{\varepsilon}{2}\right)$, 即

$$\|z - y_i\| < \frac{\varepsilon}{2}.$$

由此, 有 $\|y - y_i\| \leqslant \|y - z\| + \|z - y_i\| < \varepsilon$, 即 $y \in B(y_i; \varepsilon)$. 从而

$$\text{co}\, A \subset \bigcup_{i=1}^{m} B(y_i, \varepsilon).$$

由 $\varepsilon > 0$ 的任意性, 即得 co A 为准紧的. □

习题

3.1-1 记号与假设同定理 3.1-2. 证明: 如果空间 $(X, \|\cdot\|)$ 是可分的, 则其完备化空间 $(\widetilde{X}, \|\cdot\|_{\widetilde{X}})$ 也是可分的.

3.1-2 设 $(X, \|\cdot\|)$ 是一致凸 Banach 空间 (一致凸空间的定义见 2.17 节), Z 是 X 的非空闭的凸子集. 证明对任意给定的点 $x \in X$, 存在唯一的点 $Px \in Z$, 使得 $\|x - Px\| = \inf_{z\in Z} \|x - z\|$.

注 在 Hilbert 空间 (特殊的一致凸 Banach 空间) 中, 这个结果是基本的投影定理的一部分 (定理 4.3-1(a)).

3.2 Banach 空间的例子; 空间 $\mathcal{C}(K;Y)$, 其中 K 为紧集, Y 完备, 和空间 $\mathcal{L}(X;Y)$, 其中 Y 完备

我们从最简单的 Banach 空间的例子开始.

定理 3.2-1 任何有限维赋范向量空间是 Banach 空间.

证明 设 $(X, \|\cdot\|)$ 是 $\mathbb{K}$ 上的有限维赋范向量空间, $(e_i)_{i=1}^{n}$ 是 X 的一个基. 在 X 上赋以范数

$$\|\cdot\|_1 : x = \sum_{i=1}^{n} x_i e_i \to \|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

对任何由向量 $x^k = \sum_{i=1}^{n} x_i^k e_i \in X$ 组成的 Cauchy 序列 $(x^k)_{k=1}^{\infty}$, 有

$$\sum_{i=1}^{n} |x_i^k - x_i^\ell| = \|x^k - x^\ell\|_1$$

对任意 $k, \ell \geqslant 1$ 成立, 而数域 $\mathbb{K}$ 是完备的, 易知 X 是完备的.

因为有限维空间上所有的范数均等价 (定理 2.7-1), 所以 $(X, \|\cdot\|)$ 是完备的. □

下面的例子是基本的. 回顾一下记号 $\mathcal{C}(X;Y)$ 或 $Y = \mathbb{R}$ 时简记为 $\mathcal{C}(X)$, 表示由拓扑空间 X 到拓扑空间 Y 的所有连续映射的集合.

定理 3.2-2 设 K 为紧拓扑空间, $(Y;\|\cdot\|)$ 为 Banach 空间. 在空间 $\mathcal{C}(K;Y)$ 上赋以 sup 范数 $|||\cdot|||$, 即对任一 $f\in\mathcal{C}(K,Y)$ 定义

$$|||f|||:=\sup_{x\in K}\|f(x)\|$$

(定理 2.3-1). 则 $\mathcal{C}(K;Y)$ 为 Banach 空间.

证明 设 $(f_n)_{n=1}^{\infty}$ 是空间 $(\mathcal{C}(K;Y),|||\cdot|||)$ 上的 Cauchy 序列. 对任意给定的 $x\in K$, 当 $m,n\geqslant 1$ 时

$$\|f_m(x)-f_n(x)\|\leqslant|||f_m-f_n|||,$$

故 $(f_n(x))_{n=1}^{\infty}$ 是完备空间 $(Y,\|\cdot\|)$ 中的 Cauchy 序列, 从而收敛. 定义映射 $f:K\to Y$ 为

$$f(x):=\lim_{n\to\infty}f_n(x),\quad x\in K.$$

对任意给定的 $\varepsilon>0$, 存在 $n_0=n_0(\varepsilon)\geqslant 1$, 使得当 $m,n\geqslant n_0$ 时, 对任何 $x\in K$,

$$\|f_m(x)-f_n(x)\|\leqslant|||f_m-f_n|||\leqslant\varepsilon.$$

固定 $x\in K$, 令 $m\to\infty$, 即得当 $n\geqslant n_0$ 时

$$\|f(x)-f_n(x)\|\leqslant\varepsilon.$$

上式对任何 $x\in K$ 成立, 因此又等价于当 $n\geqslant n_0$ 时

$$\sup_{x\in K}\|f(x)-f_n(x)\|\leqslant\varepsilon.$$

剩下只要证明映射 $f:K\to Y$ 是连续的, 根据这个性质我们就可以把上面最后一个不等式的左边表示为 $|||f-f_n|||$.

设 x_0 是 K 中任意一点. 对给定的 $\varepsilon>0$, 如上选取 $n_0=n_0(\varepsilon)$. 因为映射 $f_{n_0}:K\to Y$ 在 x_0 处连续, 从而存在 x_0 的邻域 $V(x_0)\subset K$, 使得当 $x\in V(x_0)$ 时

$$\|f_{n_0}(x)-f_{n_0}(x_0)\|\leqslant\varepsilon.$$

因而 $x\in V(x_0)$ 时

$$\|f(x)-f(x_0)\|\leqslant\|f(x)-f_{n_0}(x)\|+\|f_{n_0}(x)-f_{n_0}(x_0)\|+\|f_{n_0}(x_0)-f(x_0)\|\leqslant 3\varepsilon.$$

由 $3\varepsilon>0$ 的任意性, 即得 $f:K\to Y$ 是连续的, 证毕. □

特别地, 空间 $(\mathcal{C}(\overline{\Omega}),\|\cdot\|)$ 提供了 Banach 空间的一个基本的例子, 其中 Ω 为 $\mathbb{R}^n$ 的一个有界开子集, $\|\cdot\|$ 是对每个 $f\in\mathcal{C}(\overline{\Omega})$ 由

$$\|f\|:=\sup_{x\in\overline{\Omega}}|f(x)|$$

定义的 sup 范数.

注 (1) 对任何整数 $m \geqslant 1$ 和 $\mathbb{R}^n$ 的任何有界开子集 Ω, 由所有在 $\mathbb{R}^n$ 上 m 阶连续可微函数在 $\overline{\Omega}$ 上的限制组成的空间 $\mathcal{C}^m(\overline{\Omega})$ 提供了 Banach 空间的另一个例子 (习题 3.2-1).

(2) 设 Ω 为 $\mathbb{R}^n$ 的有界开子集, 在空间 $\mathcal{C}(\overline{\Omega})$ 上赋以任何一个范数 $\|\cdot\|_{L^p(\Omega)}$, $1 \leqslant p < \infty$, 得到的赋范空间是不完备的 (习题 3.2-2).

类似于定理 3.2-2 的证明 (故而从略), 可以得到

定理 3.2-3 设 X 为任意集合, Y 为 Banach 空间, 空间 $\mathcal{B}(X;Y)$ 由 X 到 Y 的所有有界映射组成, 对任何 $f \in \mathcal{B}(X;Y)$, 其范数定义为

$$|||f||| := \sup_{x \in X} \|f(x)\|$$

(定理 2.3-2), 则 $\mathcal{B}(X;Y)$ 为 Banach 空间.

我们用 Banach 空间另一个基本的例子结束本节. 回顾一下, 给定同一个域上的两个赋范向量空间 X 和 Y. 用 $\mathcal{L}(X;Y)$ 表示由所有连续线性算子 $A: X \to Y$ 组成的赋范向量空间, 这里 $\|A\|_{\mathcal{L}(X;Y)} = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X}$ (定理 2.9-5).

定理 3.2-4 设 X 为赋范向量空间, Y 为 Banach 空间, 则 $(\mathcal{L}(X;Y), \|\cdot\|_{\mathcal{L}(X;Y)})$ 是 Banach 空间.

特别地, 对于 $\mathbb{K}$ 上赋范向量空间 X 的对偶空间 $X' = \mathcal{L}(X;\mathbb{K})$, 赋以范数

$$x' \in X' \to \|x'\| = \sup_{x \neq 0} \frac{|x'(x)|}{\|x\|_X}$$

后为 Banach 空间.

证明 为简明起见, 用同一个记号 $\|\cdot\|$ 表示空间 X, Y 和 $\mathcal{L}(X;Y)$ 上的范数. 设 $(A_n)_{n=1}^{\infty}$ 为空间 $\mathcal{L}(X;Y)$ 中的 Cauchy 序列. 对任意的 $x \in X$ 和一切 $m, n \geqslant 1$, 由不等式

$$\|A_m x - A_n x\| \leqslant \|A_m - A_n\| \, \|x\|,$$

所以 $(A_n x)_{n=1}^{\infty}$ 是 Banach 空间 Y 中的 Cauchy 序列. 于是这个序列收敛. 定义映射 $A: X \to Y$ 为

$$Ax := \lim_{n \to \infty} A_n x, \quad x \in X.$$

因为对任何数 $\alpha, \widetilde{\alpha} \in \mathbb{K}$ 和向量 $x, \widetilde{x} \in X$,

$$\begin{aligned}
A(\alpha x + \widetilde{\alpha}\widetilde{x}) &= \lim_{n \to \infty} A_n(\alpha x + \widetilde{\alpha}\widetilde{x}) \\
&= \lim_{n \to \infty} (\alpha A_n x + \widetilde{\alpha} A_n \widetilde{x}) \\
&= \alpha \lim_{n \to \infty} A_n x + \widetilde{\alpha} \lim_{n \to \infty} A_n \widetilde{x} \\
&= \alpha A x + \widetilde{\alpha} A \widetilde{x},
\end{aligned}$$

所以 A 是线性算子 (这里用到加法和数乘的连续性).

记 $C := \sup_{n\geqslant 1}\|A_n\| < \infty$(因为 Cauchy 序列是有界的, 见定理 1.12-1(a)). 对任何 $x\in X$, 由当 $n\geqslant 1$ 时 $\|A_nx\|\leqslant C\|x\|$ 和

$$\|Ax\| = \lim_{n\to\infty}\|A_nx\|,$$

即得 $\|Ax\|\leqslant C\|x\|$. 从而线性算子 $A: X\to Y$ 是连续的.

下面证明当 $n\to\infty$ 时 $\|A_n - A\|\to 0$. 任意给定 $\varepsilon > 0$, 取 $n_0 = n_0(\varepsilon)$, 使得当 $m,n\geqslant n_0$ 时 $\|A_m - A_n\|\leqslant\varepsilon$, 因此, 当 $m,n\geqslant n_0$ 时, 对所有的 $x\in X$,

$$\|A_mx - A_nx\|\leqslant\varepsilon\|x\|.$$

在上面的不等式中, 令 $m\to\infty$, 得到当 $n\geqslant n_0$ 时, 对所有的 $x\in X$ 均有

$$\|Ax - A_nx\|\leqslant\varepsilon\|x\|.$$

因此, 当 $n\geqslant n_0$ 时,

$$\|A_n - A\| = \sup_{x\neq 0}\frac{\|A_nx - Ax\|}{\|x\|}\leqslant\varepsilon,$$

这就证得结论. □

Banach 空间的另一个重要的例子是当 X 自身为 Banach 空间时的商空间 X/Z (2.2 节) (定理 3.6-5), 其他的例子可见习题 3.2-4 和 3.2-5.

习题

3.2-1 设 Ω 为 $\mathbb{R}^n$ 中的一个区域. 证明对任何整数 $m\geqslant 1$, 空间 $\mathcal{C}^m(\overline{\Omega})$ (定理 1.18-1) 是 Banach 空间, 其中范数 $\|\cdot\|_{\mathcal{C}^m(\overline{\Omega})}$ 定义为对任何 $f\in\mathcal{C}^m(\overline{\Omega})$,

$$\|f\|_{\mathcal{C}^m(\overline{\Omega})} := \max_{|\alpha|\leqslant m}\sup_{x\in\Omega}|\partial^\alpha f(x)|.$$

3.2-2 设 Ω 为 $\mathbb{R}^n$ 中的有界开子集, $1\leqslant p<\infty$. 证明空间 $(\mathcal{C}(\overline{\Omega}), \|\cdot\|_{L^p(\Omega)})$ 不是完备的.

提示: 构造一个不收敛的 Cauchy 序列.

3.2-3 证明: 所有多项式 $p:\mathbb{R}\to\mathbb{R}$ 组成的空间 $\mathcal{P}$ 赋以范数 $\|p\| = \sup_{0\leqslant x\leqslant 1}|p(x)|$ 不是完备空间.

注 实际上, 不存在使 $\mathcal{P}$ 成为 Banach 空间的范数. 我们将会看到 (定理 5.1-4), 这个非平凡的结果是 Baire 定理 (5.1 节) 的一个推论.

3.2-4 设 X 和 Y 是同一个域上的两个赋范向量空间. 证明: 如果 Y 是完备的, 则由 X 到 Y 的所有紧算子组成的 $\mathcal{L}(X;Y)$ 的子空间 $\mathcal{K}(X;Y)$ 在 $\mathcal{L}(X;Y)$ 中闭. 因此, 当 Y 为 Banach 空间时, 作为 Banach 空间的闭子集, $\mathcal{K}(X;Y)$ 也是 Banach 空间.

3.2-5 设 $X_1, X_2,\cdots, X_k, k\geqslant 2$, 和 Y 均为赋范向量空间. 证明: 如果 Y 是完备的, 则由 $X_1\times X_2\times\cdots\times X_k$ 到 Y 的所有连续多重线性映射 (2.11 节) 组成的空间 $\mathcal{L}_k(X_1, X_2,\cdots, X_k; Y)$ 是完备的.

3.3 取值于 Banach 空间的单实变量连续函数的积分

作为连续线性延拓的唯一性 (定理 3.1-1) 和赋范向量空间完备化 (定理 3.1-2) 的一个有趣应用, 便是积分 $\int_a^b f(x)dx$ 的构造, 其中 $f:[a,b]\to Y$ 是取值于 Banach 空间 Y 的连续函数. 这类积分将用于 Banach 空间的中值定理的证明 (定理 7.6-1), 后者又将用于建立 Banach 空间中的 Newton-Kantorovich 定理 (定理 7.7-3). 而且, 积分 $\int_a^b f(x)dx$ 又自然地导致下面记作 $\mathcal{R}([a,b];Y)$ 的另一个 Banach 空间的定义, 这个空间包含空间 $\mathcal{C}([a,b];Y)$.

注 更一般地, 可以构造定义于测度空间 (1.14 节) 取值于 Banach 空间的函数的 Lebesgue 积分 (只要对这些函数适当地定义可测性的概念)[2].

$\int_a^b f(x)dx$ 的定义将分两步进行.

首先, 设 $f:[a,b]\subset\mathbb{R}\to Y$ 是 $[a,b]$ 上的阶梯函数. 即存在有限个点 $x_i\in[a,b], 1\leqslant i\leqslant n$, 和向量 $c_i\in Y, 1\leqslant i\leqslant n$, 使得

$$a=x_0<x_1<\cdots<x_i<\cdots<x_{n-1}<x_n=b,$$
$$f(x)=c_i,\ \text{其中}\ x_{i-1}<x<x_i,\ 1\leqslant i\leqslant n,$$
$$\max_{0\leqslant i\leqslant n}\|f(x_i)\|_Y\leqslant\max_{1\leqslant i\leqslant n}\|c_i\|_Y.$$

自然地定义这类阶梯函数的积分为

$$\ell(f):=\sum_{i=1}^n(x_i-x_{i-1})c_i\in Y.$$

易见所有取值于 Y 的 $[a,b]$ 上的阶梯函数的集合 $\mathcal{S}([a,b];Y)$ 是一个向量空间, 又赋予定义为

$$\|f\|:=\sup_{a\leqslant x\leqslant b}\|f(x)\|_Y$$

的 sup 范数后, $\mathcal{S}([a,b];Y)$ 成为赋范向量空间. 上述映射 $\ell:\mathcal{S}([a,b];Y)\to Y$ 显然是线性的, 而且, 由 $\ell(f)$ 的定义直接可得对任何 $f\in\mathcal{S}([a,b];Y)$,

$$\|\ell(f)\|_Y\leqslant(b-a)\|f\|,$$

从而 ℓ 在这个空间上连续.

其次, 用 $\mathcal{R}([a,b];Y)$ 表示空间 $\mathcal{S}([a,b];Y)$ 关于 sup 范数 $\|\cdot\|$ 的完备化空间. Banach 空间 $\mathcal{R}([a,b];Y)$ 是赋以 sup 范数 (定理 3.2-3) 的从 $[a,b]$ 到 Y 的所有有界函数组成的 Banach 空间 $\mathcal{B}([a,b];Y)$ 的闭子空间. 由构造, $\mathcal{S}([a,b];Y)$ 在 $\mathcal{R}([a,b];Y)$ 中稠密, 而且 Y 是完备的, 所以连续线性映射 $\ell:\mathcal{S}([a,b];Y)\to Y$ 可以唯一地连续延拓

---

2) 例如可见 SCHWARTZ [1993a] 或 LANG [1993].

到空间 $\mathcal{R}([a,b];Y)$ 上 (定理 3.1-1 和 3.1-2), 这也是要求 Y 完备性的本质所在. 由此, 可以导出对任何函数 $f \in \mathcal{R}([a,b];Y)$ 在 $[a,b]$ 上积分的自然定义

$$\int_a^b f(x)dx := \lim_{n\to\infty} \ell(f_n),$$

这里 $(f_n)_{n=1}^\infty$ 是阶梯函数 $f_n \in \mathcal{S}([a,b];Y)$ 的序列, 满足当 $n \to \infty$ 时 $\|f_n - f\| \to 0$, 其中范数 $\|\cdot\|$ 的定义满足对每个 $f \in \mathcal{R}([a,b];Y)$,

$$\|f\| = \sup_{a\leqslant x\leqslant b} \|f(x)\|_Y.$$

这样构造的 Banach 空间 $\mathcal{R}([a,b];Y)$ 和积分 $\int_a^b f(x)dx$ 具有下述性质.

定理 3.3-1 (a) 对每个函数 $f \in \mathcal{R}([a,b];Y)$,

$$\left\|\int_a^b f(x)dx\right\|_Y \leqslant \int_a^b \|f(x)\|_Y dx \leqslant (b-a)\|f\|.$$

(b) 空间 $\mathcal{R}([a,b];Y)$ 包含空间 $\mathcal{C}([a,b];Y)$.

证明 显然, 对任何阶梯函数 $f \in \mathcal{S}([a,b];Y)$, 成立不等式

$$\|\ell(f)\|_Y \leqslant \int_a^b \|f(x)\|_Y dx \leqslant (b-a)\|f\|.$$

这个不等式中的每一项均是 $f \in \mathcal{S}([a,b];Y)$ 的连续函数, 从而不等式对 $\mathcal{S}([a,b];Y)$ 的闭包 $\mathcal{R}([a,b];Y)$ 中的任何函数均成立.

再设 $f : [a,b] \to Y$ 为任一连续函数. 由区间 $[a,b]$ 的紧性, f 在 $[a,b]$ 上一致连续, 由此可知 f 是一列阶梯函数 $f_n \in \mathcal{S}([a,b];Y), n \geqslant 1$ 的一致收敛的极限. 作为极限 $\lim_{n\to\infty} \ell(f_n)$ 它与逼近于 f 的阶梯函数序列的选取无关, 因此, $\int_a^b f(x)dx \in Y$ 是一意确定的. □

注 空间 $\mathcal{R}([a,b];Y)$ 中的函数称之为*正则函数*. 由上述讨论可见, 由所有取值于 Y 的连续函数组成的空间 $\mathcal{C}([a,b];Y)$ 满足包含关系 $\mathcal{C}([a,b];Y) \subset \mathcal{R}([a,b];Y)$. 而且, 可以证明这个包含关系还是严格的. 例如, 空间 $\mathcal{R}([a,b];Y)$ 包含所有的单调函数, 这类函数可能在 $[a,b]$ 中的可数个点处不连续[3].

3.4 Banach 空间的例: 空间 ℓ^p 和 $L^p(\Omega)$, $1 \leqslant p \leqslant \infty$

我们由考察在 2.4 节中介绍的赋范向量空间 $(\ell^p, \|\cdot\|_p)$, $1 \leqslant p \leqslant \infty$, 开始. 数量 x_i 的序列 $x = (x_i)_{i=1}^\infty \in \ell^p$ 的范数定义为

[3] 关于这些概念的更详细的讨论可参见 DIEUDONNÉ [1960] 或 LANG [1993].

$$\|x\|_p := \left(\sum_{i=1}^{\infty}|x_i|^p\right)^{\frac{1}{p}}, \quad 1 \leqslant p < \infty,$$
$$\|x\|_\infty := \sup_{i\geqslant 1}|x_i|, \quad p = \infty.$$

定理 3.4-1 空间 $(\ell^p, \|\cdot\|_p), 1 \leqslant p \leqslant \infty$, 为 Banach 空间.

证明 设 $(x^n)_{n=1}^\infty$ 为元素 $x^n = (x_i^n)_{i=1}^\infty \in \ell^p$ 的 Cauchy 序列. 因为对每个 $i \geqslant 1$,

$$|x_i^m - x_i^n| \leqslant \|x^m - x^n\|_p,$$

故而每个数列 $(x_i^n)_{n=1}^\infty$ 是 Canuchy 数列. 令

$$x := (x_i)_{i=1}^\infty,$$

这里对每个 $i \geqslant 1, x_i := \lim_{n\to\infty} x_i^n$.

首先, 我们证明 $x \in \ell^p$. 设 M 满足对所有的 $n \geqslant 1$, $\|x^n\|_p \leqslant M$ (Cauchy 序列是有界的). 因此, 对任何整数 $k \geqslant 1$,

$$\left(\sum_{i=1}^{k}|x_i|^p\right)^{\frac{1}{p}} = \lim_{n\to\infty}\left(\sum_{i=1}^{k}|x_i^n|^p\right)^{\frac{1}{p}} \leqslant M, \quad 1 \leqslant p < \infty,$$
$$\sup_{1\leqslant i\leqslant k}|x_i| = \lim_{n\to\infty}\sup_{1\leqslant i\leqslant k}|x_i^n| \leqslant M, \quad p = \infty.$$

因为上界 M 与整数 k 无关, 在上面两式中令左边的 $k \to \infty$, 即得 $x \in \ell^p$.

其次, 说明当 $n \to \infty$ 时 $\|x^n - x\|_p \to 0$. 对任意给定的 $\varepsilon > 0$, 因为 $(x^n)_{n=1}^\infty$ 是 ℓ^p 中 Cauchy 序列, 故存在 $n_0 = n_0(\varepsilon)$, 使当 $m, n \geqslant n_0$ 时

$$\left(\sum_{i=1}^{\infty}|x_i^m - x_i^n|^p\right)^{\frac{1}{p}} \leqslant \varepsilon, \quad 1 \leqslant p < \infty$$

或

$$\sup_{i\geqslant 1}|x_i^m - x_i^n| \leqslant \varepsilon, \quad p = \infty,$$

由此, 对任何整数 $k \geqslant 1$, 当 $m, n \geqslant n_0$ 时

$$\left(\sum_{i=1}^{k}|x_i^m - x_i^n|^p\right)^{\frac{1}{p}} \leqslant \varepsilon, \quad 1 \leqslant p < \infty,$$
$$\sup_{1\leqslant i\leqslant k}|x_i^m - x_i^n| \leqslant \varepsilon, \quad p = \infty.$$

固定 k, 令 $m \to \infty$, 即得当 $n \geqslant n_0$ 时

$$\left(\sum_{i=1}^{k}|x_i - x_i^n|^p\right)^{\frac{1}{p}} \leqslant \varepsilon, \quad 1 \leqslant p < \infty,$$
$$\sup_{1\leqslant i\leqslant k}|x_i - x_i^n| \leqslant \varepsilon, \quad p = \infty,$$

令左边的 $k \to \infty$, 即得当 $n \geqslant n_0$ 时

$$\|x - x^n\|_p \leqslant \varepsilon.$$

这就得到了结论. □

现在我们来考察在 2.5 节中介绍过的 (实) Lebesgue 空间 $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)}), 1 \leqslant p \leqslant \infty$, 其中 Ω 是 $\mathbb{R}^n$ 中的任意开子集. 前面已经说明函数 $f \in L^p(\Omega)$ 在 Ω 上几乎处处有限, 其范数定义为

$$\|f\|_{L^p(\Omega)} = \left(\int_\Omega |f(x)|^p dx\right)^{\frac{1}{p}}, \quad 1 \leqslant p < \infty,$$
$$\|f\|_{L^\infty(\Omega)} = \inf\{C \geqslant 0;\ \text{在 } \Omega \text{ 几乎处处 } |f| \leqslant C\},\ p = \infty.$$

不难理解, 建立这些空间的完备性比空间 ℓ^p 稍复杂些.

定理 3.4-2 空间 $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$, $1 \leqslant p \leqslant \infty$ 是 Banach 空间, 其中 Ω 是 $\mathbb{R}^n$ 的开子集.

证明 为简洁起见, 在证明中记 $\|\cdot\|_p = \|\cdot\|_{L^p(\Omega)}$. 我们从考察 $1 \leqslant p < \infty$ 的情况开始. 设 $(f_m)_{m=1}^\infty$ 是函数 $f_m \in L^p(\Omega)$ 的 Cauchy 序列. 于是, 存在子列 $(f_{\sigma(m)})_{m=1}^\infty$ 满足当 $m \geqslant 1$ 时

$$\|f_{\sigma(m+1)} - f_{\sigma(m)}\|_p \leqslant \frac{1}{2^m}.$$

对每个整数 $k \geqslant 1$, 定义函数 g_k 为

$$g_k := \sum_{m=1}^k |f_{\sigma(m+1)} - f_{\sigma(m)}|,$$

显然 g_k 属于空间 $L^p(\Omega)$, 而且满足

$$0 \leqslant g_1 \leqslant \cdots \leqslant g_k \leqslant g_{k+1} \leqslant \cdots,$$
$$\|g_k\|_p \leqslant \sum_{m=1}^k \frac{1}{2^m} \leqslant 1,$$

因此, 对所有的 $x \in \Omega$

$$g(x) := \lim_{k\to\infty} g_k(x)$$

在 $[0, \infty]$ 中存在.

将 Fatou 引理 (定理 1.15-2) 应用于如上定义的函数 $g : \Omega \to [0, \infty]$, 可得

$$\int_\Omega |g(x)|^p dx = \int_\Omega \lim_{k\to\infty} |g_k(x)|^p dx$$
$$\leqslant \liminf_{k\to\infty} \int_\Omega |g_k(x)|^p dx \leqslant 1.$$

因此, $g \in L^p(\Omega)$, 且对几乎所有的 $x \in \Omega$, $0 \leqslant g(x) < \infty$. 因为对所有的 $x \in \Omega$,

$$\sum_{m=1}^{k} |f_{\sigma(m+1)}(x) - f_{\sigma(m)}(x)| = g_k(x) \leqslant g(x),$$

所以对几乎所有的 $x \in \Omega$,

$$\begin{aligned} f(x) :&= \lim_{k\to\infty} f_{\sigma(k+1)}(x) \\ &= \lim_{k\to\infty} \left(f_{\sigma(1)}(x) + \sum_{m=1}^{k} (f_{\sigma(m+1)}(x) - f_{\sigma(m)}(x)) \right) \end{aligned}$$

在 $\mathbb{R}$ 中存在, 另一方面, 这样定义的函数 f 属于 $L^p(\Omega)$. 这是因为对几乎所有的 $x \in \Omega$,

$$\begin{aligned} |f(x)| &\leqslant |f_{\sigma(1)}(x)| + \sum_{m=1}^{k} |f_{\sigma(m+1)}(x) - f_{\sigma(m)}(x)| \\ &= |f_{\sigma(1)}(x)| + g_k(x) \leqslant |f_{\sigma(1)}(x)| + g(x), \end{aligned}$$

而函数 $f_{\sigma(1)}$ 和 g 均在 $L^p(\Omega)$ 中 (注意 $g \geqslant 0$).

还需证明当 $m \to \infty$ 时 $\|f_m - f\|_p \to 0$. 给定 $\varepsilon > 0$, 取 $m_0 = m_0(\varepsilon)$, 使得当 ℓ, $m \geqslant m_0$ 时对所有 $\|f_\ell - f_m\|_p \leqslant \varepsilon$. 再次应用 Fatou 引理, 可知当 $m \geqslant m_0$ 时

$$\begin{aligned} \int_\Omega |f(x) - f_m(x)|^p dx &= \int_\Omega \lim_{k\to\infty} |f_{\sigma(k)}(x) - f_m(x)|^p dx \\ &\leqslant \liminf_{k\to\infty} \int_\Omega |f_{\sigma(k)}(x) - f_m(x)|^p dx \leqslant \varepsilon^p, \end{aligned}$$

即当 $m \geqslant m_0$ 时 $\|f - f_m\|_p \leqslant \varepsilon$. 这就对 $1 \leqslant p < \infty$ 证得结论.

给定 $L^\infty(\Omega)$ 中的 Cauchy 序列 $(f^m)_{m=1}^\infty$. 设 M 满足对所有的 $m \geqslant 1$ 均有 $\|f^m\|_\infty \leqslant M$, 这样, 对几乎所有的 $x \in \Omega$,

$$\begin{aligned} |f^m(x)| &\leqslant M, \\ |f^\ell(x) - f^m(x)| &\leqslant \|f^\ell - f^m\|. \end{aligned}$$

类同于定理 3.2-2 的推理, 可知对几乎所有的 $x \in \Omega$, $f(x) := \lim_{m\to\infty} f_m(x)$ 存在, 且这样定义的函数 f 在 $L^\infty(\Omega)$ 中, 当 $m \to \infty$ 时

$$\|f^m - f\|_\infty \to 0. \qquad \square$$

上述证明的第一部分还导出了 $L^p(\Omega)$ 中收敛序列的一个值得注意的性质 (因为已经证明了 $L^p(\Omega)$ 中的收敛序列与 Cauchy 序列是一致的).

定理 3.4-3 设 $(f_m)_{m=1}^\infty$ 是 $L^p(\Omega)$ 中的收敛序列, 这里 $1 \leqslant p < \infty$, 其极限为 $f \in L^p(\Omega)$, 则存在子列 $(f_{\sigma(m)})_{m=1}^\infty$ 在 Ω 上几乎处处收敛于 f, 即对几乎所有的 $x \in \Omega$,

$$\lim_{m\to\infty} f_{\sigma(m)}(x) = f(x).$$

习题

3.4-1 证明对 $0 < p < 1$, 习题 2.5-4 中定义的距离空间 $(L^p(\Omega), d_p)$ 是完备的.

3.5 赋范向量空间的对偶; 例; $L^p(\Omega)(1 \leqslant p < \infty)$ 中的 F. Riesz 表示定理

设 X 为域 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的赋范向量空间. 前面已引入空间

$$X' := \mathcal{L}(X, \mathbb{K}),$$

称之为 X 的对偶空间, 或简称为 X 的对偶 (2.9 节), 它由所有在 X 上连续的线性泛函 $x' : X \to \mathbb{K}$ 组成. 因为 $\mathbb{K}$ 是完备的, 所以在 X' 上赋以算子范数, 即对 $x' \in X'$, 定义

$$\|x'\| = \sup_{x \neq 0} \frac{|x'(x)|}{\|x\|}$$

所得的赋范向量空间为 Banach 空间 (定理 3.2-4), 这一性质和 X 完备与否无关.

对偶空间在线性泛函分析中起着核心的作用, 正如将在第 5 章中详细说明的, 它们的基本性质将会得到系统的研究. 本节的目的则是给出对偶空间的几个基本的例子.

任意给定扩充实数 $1 \leqslant p \leqslant \infty$, 定义对应的扩充实数 $1 \leqslant q \leqslant \infty$ 为

当 $p = 1$ 时, $q = \infty$,

当 $1 < p < \infty$ 时, $\frac{1}{p} + \frac{1}{q} = 1$,

当 $p = \infty$ 时, $q = 1$,

称 q 为 p 的共轭指数.

作为开始, 我们先考察在第 2.4 节中介绍的空间 $(\ell^p, \|\cdot\|_p), 1 \leqslant p \leqslant \infty$. 在下面的定理中将会看到, 值得注意的是当 $1 \leqslant p < \infty$ 时 ℓ^p 的对偶可恒同于空间 ℓ^q 其中 q 表示 p 的共轭指数. 要注意当 $p = \infty$ 时, 这个结论并不成立 (定理 3.5-2).

定理 3.5-1 (ℓ^p 的对偶, $1 \leqslant p < \infty$) 给定实数 $1 \leqslant p < \infty$, 用 $1 < q \leqslant \infty$ 表示共轭指数. 则对任意给定的元素 $a = (a_i)_{i=1}^{\infty} \in \ell^q$, 由

$$x'(x) = \sum_{i=1}^{\infty} a_i x_i, \quad x = (x_i)_{i=1}^{\infty} \in \ell^p$$

定义了 ℓ^p 上的连续线性泛函 x', 而且

$$\|x'\|_{(\ell^p)'} = \|a\|_q.$$

这样定义的线性等距算子 $a \in \ell^q \to x' \in (\ell^p)'$ 是一个双射, 即对 ℓ^p 上任意给定的连续线性泛函 x', 存在唯一的元素 $a = (a_i)_{i=1}^{\infty} \in \ell^q$, 使得对任意的 $x = (x_i)_{i=1}^{\infty} \in \ell^p$, 有 $x'(x) = \sum_{i=1}^{\infty} a_i x_i$.

因此, 对任意的 $1 \leqslant p < \infty$, ℓ^p 的对偶空间, 作为一个赋范向量空间, 恒同于空间 ℓ^q.

证明 (i) 由序列的 Hölder 不等式 (定理 2.4-1), 当 $1 < p < \infty$ 时, 对所有的 $a = (a_i)_{i=1}^{\infty} \in \ell^q$ 和 $x = (x_i)_{i=1}^{\infty} \in \ell^p$, 成立

$$\left|\sum_{i=1}^{\infty} a_i x_i\right| \leqslant \|a\|_q \|x\|_p,$$

这个不等式对 $p = 1, q = \infty$ 显然成立. 这样, 给定 $a = (a_i)_{i=1}^{\infty} \in \ell^q$, 对所有的 $x = (x_i)_{i=1}^{\infty} \in \ell^p$, 令 $x'(x) = \sum\limits_{i=1}^{\infty} a_i x_i$ 就定义了 ℓ^p 上的一个连续线性泛函 x', 且其范数满足 $\|x'\| \leqslant \|a\|_q$.

为建立相反的不等式, 分两种情况讨论. 首先, 设 $p = 1$. 对任意整数 $n \geqslant 1$, 令

$$x^n = (x_i^n)_{i=1}^{\infty} \in \ell^p,$$

其中 $x_n^n := \mathrm{sgn}\ a_n$, 当 $i \neq n$ 时 $x_i^n := 0$, 这样, $x'(x^n) = a_n x_n^n = |a_n|$, 且 $\|x^n\|_1 \leqslant 1$ (因为当 $a_n \neq 0$ 时 $\|x^n\|_1 = 1$, 当 $a_n = 0$ 时, $\|x^n\|_1 = 0$), 从而

$$|a_n| = |x'(x^n)| \leqslant \|x'\|\ \|x^n\|_1 \leqslant \|x'\|,$$

由此可得 $\|a\|_\infty = \sup_{n \geqslant 1} |a_n| \leqslant \|x'\|$. 所以, 当 $p = 1$ 时 $\|x'\| = \|a\|_\infty$.

其次, 设 $1 < p < \infty$, 对任何整数 $n \geqslant 1$, 令

$$x^n = (x_i^n)_{i=1}^{\infty},$$

其中当 $1 \leqslant i \leqslant n$ 时, $x_i^n := |a_i|^{q-1}\mathrm{sgn}\ a_i$, 当 $i > n$ 时 $x_i^n = 0$. 于是, $x'(x^n) = \sum\limits_{i=1}^{n} a_i x_i = \sum\limits_{i=1}^{n} |a_i|^q, \|x^n\|_p = \left(\sum\limits_{i=1}^{n} |a_i|^q\right)^{\frac{1}{p}}$, 从而

$$\sum_{i=1}^{n} |a_i|^q = |x'(x^n)| \leqslant \|x'\|\ \|x^n\|_p = \|x'\| \left(\sum_{i=1}^{n} |a_i|^q\right)^{\frac{1}{p}}.$$

由此可知 $\|a\|_q = \lim_{n\to\infty} \left(\sum\limits_{i=1}^{n} |a_i|^q\right)^{\frac{1}{q}} \leqslant \|x'\|$. 因此, 当 $1 < p < \infty$ 时, $\|x'\| = \|a\|_p$.

(ii) 余下要证的是对任何 $1 \leqslant p < \infty$, 在 (i) 中定义的等距算子 $a \in \ell^q \to x' \in (\ell^p)'$ 是满射 (显然它是线性的单射). 设给定 $x' \in (\ell^p)'$.

对任意整数 $i \geqslant 1$, 定义元素 $e_i \in \ell^p$ 和数 $a_i \in \mathbb{K}$ 为

$$e_i := (\delta_{ij})_{j=1}^{\infty}, \quad a_i := x'(e_i).$$

给定任意的 $x = (x_i)_{i=1}^{\infty} \in \ell^p$, 由关系式 $\|x - \sum_{i=1}^{n} x_i e_i\|_p = \left(\sum_{i=n+1}^{\infty} |x_i|^p\right)^{\frac{1}{p}}$, 可知在空间 ℓ^p 中, $\lim_{n\to\infty} \sum_{i=1}^{n} x_i e_i = x$. 由 x' 的连续性可知, 当 $n \to \infty$ 时

$$x'\left(\sum_{i=1}^{n} x_i e_i\right) = \sum_{i=1}^{n} a_i x_i \to x'(x).$$

因此, 级数 $\sum_{i=1}^{\infty} a_i x_i$ 在 $\mathbb{K}$ 中收敛, 其和即 $x'(x)$. □

再考察 $p = \infty$ 的情况.

定理 3.5-2 任意给定元素 $a = (a_i)_{i=1}^{\infty} \in \ell^1$, 由关系式

$$x'(x) = \sum_{i=1}^{\infty} a_i x_i,$$

其中 $x = (x_i)_{i=1}^{\infty} \in \ell^{\infty}$, 定义了 ℓ^{∞} 上的一个连续线性泛函 x', 而且

$$\|x'\|_{(\ell^{\infty})'} = \|a\|_1.$$

这样定义的等距算子 $a \in \ell^1 \to x' \in (\ell^{\infty})'$ 是线性单射, 但不是满射, 即赋范向量空间 ℓ^1 只恒同于 ℓ^{∞} 的对偶空间的一个真子空间.

证明 因为对任何 $a = (a_i)_{i=1}^{\infty} \in \ell^1$ 和 $x = (x_i)_{i=1}^{\infty} \in \ell^{\infty}$,

$$\left| \sum_{i=1}^{\infty} a_i x_i \right| \leqslant \|a\|_1 \|x\|_{\infty},$$

由对每个 $x = (x_i)_{i=1}^{\infty} \in \ell^{\infty}$, 关系式 $x'(x) = \sum_{i=1}^{\infty} a_i x_i$ 定义了 ℓ^{∞} 上的一个连续线性泛函 x', 其范数满足 $\|x'\| \leqslant \|a\|_1$.

为建立相反方向的不等式, 设 $x = (\operatorname{sgn} a_i)_{i=1}^{\infty}$. 于是, $x'(x) = \sum_{i=1}^{\infty} |a_i|$, 且 $\|x\|_{\infty} \leqslant 1$. 这样

$$\|a\|_1 = \sum_{i=1}^{\infty} |a_i| = |x'(x)| \leqslant \|x'\| \, \|x\|_{\infty} \leqslant \|x'\|.$$

因此, $\|x'\| = \|a\|_1$. 这样定义的显然为线性的映射 $a \in \ell^1 \to x' \in (\ell^{\infty})'$ 是等距算子 (因此是单射). □

证明这个等距算子或一般地, 任何一个从 ℓ^1 到 $(\ell^{\infty})'$ 的线性等距算子都不是满射, 最便捷的途径是借助于下述结果 (独立于现在的讨论, 由于其证明依赖 Hahn-Banach 定理, 见定理 5.9-5, 故将在以后给出): 如果赋范向量空间 X 的对偶空间是可分的, 则 X 是可分的. 这样, 如果 $(\ell^{\infty})'$ 可通过一个线性等距算子恒同于赋范向量空间 ℓ^1, 因可分性是只与范数有关的性质, 所以 $(\ell^{\infty})'$ 就与 ℓ^1 一样是可分的, 于是, ℓ^{∞} 应是可分的, 但实际上却非如此 (定理 2.4-2(c)).

现在我们来讨论在 2.5 节中介绍的 (实) Lebesgue 空间 $(L^p(\Omega), \|\cdot\|_{L^p(\Omega)})$, $1 \leqslant p \leqslant \infty$. 尽管相应的结论类似于空间 $\ell^p, 1 \leqslant p \leqslant \infty$(比较定理 3.5-1 与 3.5-2 和定理 3.5-3 与 3.5-4), 证明却要复杂得多. 注意, 在本节下面部分, 为表示 $L^p(\Omega)$ 的对偶空间的一般元素, 我们将采用记号 ℓ 而不用 x' (因为通常 x 将表示集合 Ω 中一般的点, 这样可避免误解). 下面的结果是基本的.

定理 3.5-3 (**$L^p(\Omega)$ 中的 F. Riesz 表示定理, $1 \leqslant p < \infty$**[4]) 设 Ω 为 $\mathbb{R}^n$ 中的开子集, 给定实数 $1 \leqslant p < \infty$, 用 $1 < q \leqslant \infty$ 表示 p 的共轭指数. 则对任意给定的 $g \in L^q(\Omega)$, 由对所有的 $f \in L^p(\Omega)$,

$$\ell(f) = \int_\Omega f(x)g(x)dx$$

在 $L^p(\Omega)$ 上定义了一个连续线性泛函 ℓ, 而且

$$\|\ell\|_{(L^p(\Omega))'} = \|g\|_{L^q(\Omega)}.$$

这样定义的线性等距算子 $g \in L^q(\Omega) \to \ell \in (L^p(\Omega))'$ 是一个双射, 即对任意给定的 $L^p(\Omega)$ 上的连续线性泛函 ℓ, 存在唯一的函数 $g \in L^q(\Omega)$, 使得对任何 $f \in L^p(\Omega)$, $\ell(f) = \int_\Omega f(x)g(x)dx$.

因此, $L^p(\Omega)$ 的对偶空间, 这里 $1 \leqslant p < \infty$, 作为一个赋范向量空间恒同于空间 $L^q(\Omega)$.

证明 为记号简洁起见, 在证明中范数 $\|\cdot\|_{L^p(\Omega)}$ 或 $\|\cdot\|_{(L^q(\Omega))'}$ 将被简记为 $\|\cdot\|_{L^p}$ 或 $\|\cdot\|_{(L^q)'}$.

(i) 如果 $1 < p < \infty$, 由函数的 Hölder 不等式 (定理 2.5-1), 对所有的 $f \in L^p(\Omega)$, $g \in L^q(\Omega)$, 有

$$\left|\int_\Omega fgdx\right| \leqslant \|f\|_{L^p}\|g\|_{L^q},$$

而对 $p = 1, q = \infty$, 这个不等式显然成立. 给定 $g \in L^q(\Omega)$. 对所有的 $f \in L^p(\Omega)$, 令 $\ell(f) = \int_\Omega fgdx$, 这就定义了 $L^p(\Omega)$ 上的一个连续线性泛函, 其范数满足 $\|\ell\|_{(L^p)'} \leqslant \|g\|_{L^q}$.

余下要证明这样定义的连续线性算子 $g \in L^q(\Omega) \to \ell \in (L^p(\Omega))'$ 是等距的满射. 为此, 先设 $\mu(\Omega) < \infty$, 其中 μ 表示 $\mathbb{R}^n$ 上的 Lebesgue 测度, 参见 (ii) 到 (vi).

在下面的证明中, 用 ℓ 表示 $L^p(\Omega)$ 上给定的连续线性泛函.

(ii) 设 $\mu(\Omega) < \infty$, 则存在函数 $g \in L^1(\Omega)$, 使得对任何可测的简单函数 $s : \Omega \to \mathbb{R}$,

$$\ell(s) = \int_\Omega sgdx$$

成立.

用 $\mathcal{A}$ 表示 Ω 的所有 Lebesgue 可测子集组成的 σ 代数. 因为 $\mu(\Omega) < \infty$, 所以任何 $A \in \mathcal{A}$ 的特征函数 χ_A 在 $L^p(\Omega)$ 中. 定义函数 $\nu : \mathcal{A} \to \mathbb{R}$ 为对任何 $A \in \mathcal{A}$,

$$\nu(A) := \ell(\chi_A).$$

[4] 这个命名是为了纪念 F. Riesz, 他于 1910 年开始研究了空间 ℓ^p 和 $L^p(\Omega)$, 并于 1913 年证明了这个表示定理 (对 Ω 为 $\mathbb{R}$ 中的开区间), 关于这一思想的来源在 Dieudonné [1981, 第 6 章第 2 节] 中有漂亮的阐述.

我们来证明 ν 是广义测度, 它关于 Lebesgue 测度 μ 绝对连续 (1.15 节).

首先, 由 $\chi_\varnothing = 0$, 显然可得 $\nu(\varnothing) = 0$, 而且 ν 是有限可加的, 这是因为如果 $A_1 \cap A_2 = \varnothing$, 则 $\chi_{A_1 \cup A_2} = \chi_{A_1} + \chi_{A_2}$, 由 ℓ 的线性可得 $\nu(A_1 + A_2) = \nu(A_1) + \nu(A_2)$.

其次, 对 $i \geqslant 1$, 给定可数无限个 $A_i \in \mathcal{A}$, 设它们两两互不相交, 令

$$A := \bigcup_{i=1}^{\infty} A_i, \quad B_m := A - \bigcup_{i=1}^{m} A_i,$$

其中 $m \geqslant 1$. 由 ν 的有限可加性, 对任何 $m \geqslant 1$,

$$\nu(A) = \nu\left(\bigcup_{i=1}^{m} A_i\right) + \nu(B_m) = \sum_{i=1}^{m} \nu(A_i) + \nu(B_m).$$

因为 Lebesgue 测度 μ 是可列可加的, 故而

$$\mu(A) = \sum_{i=1}^{m} \mu(A_i) + \mu(B_m) = \sum_{i=1}^{\infty} \mu(A_i) < \infty,$$

从而当 $m \to \infty$ 时 $\mu(B_m) \to 0$. 由 ℓ 的连续性, 对一切 $m \geqslant 1$,

$$|\nu(B_m)| = |\ell(\chi_{B_m})| \leqslant \|\ell\| \, \|\chi_{B_m}\|_{L^p} = \|\ell\|(\mu(B_m))^{\frac{1}{p}},$$

从而 ν 是可列可加的. 另一方面又因对任意 $A \in \mathcal{A}, |\nu(A)| = |\ell(\chi_A)| \leqslant \|\ell\|(\mu(A))^{\frac{1}{p}}$, 所以对任何 $A \in \mathcal{A}$,

$$|\nu(A)| < \infty, \text{ 且当 } \mu(A) = 0 \text{ 时 } \nu(A) = 0.$$

由 Radon-Nikodym 定理 (定理 1.15-4), 存在函数 $g \in L^1(\Omega)$, 使得对任何 $A \in \mathcal{A}$,

$$\nu(A) = \int_A g dx$$

或等价地, 对任何 $A \in \mathcal{A}$, 有

$$\ell(\chi_A) = \int_\Omega \chi_A g dx.$$

因为任何可测的简单函数形如 $s = \sum_{i=1}^{m} \alpha_i \chi_{A_i}$, 其中 $\alpha_i \in \mathbb{R}$, 而 $A_i \in \mathcal{A}$, $1 \leqslant i \leqslant m$, 由 ℓ 的线性即得对这样的函数 s, 成立 $\ell(s) = \int_\Omega s g dx$.

(iii) 设 $\mu(\Omega) < \infty$, 又设 $g \in L^1(\Omega)$ 为 (ii) 中出现的函数. 对每个整数 $k \geqslant 1$, 定义可测集

$$B_k := \{x \in \Omega; |g(x)| \leqslant k\}.$$

则对所有的 $f \in L^p(\Omega)$, 当 $f|_{\Omega - B_k} = 0$ 时,

$$\ell(f) = \int_\Omega f g dx.$$

下面, 固定 $k \geqslant 1$, 给定的 $f \in L^p(\Omega)$, 满足 $f|_{\Omega - B_k} = 0$. 由定理 1.14-5, 存在可测的简单函数 $s_m : m \geqslant 1$, 使得对几乎所有的 $x \in \Omega$,

$$|s_m(x)| \leqslant |f(x)|, \quad s_m(x) \to f(x).$$

因此, 对几乎所有的 $x \in \Omega$,

$$(s_m(x) - f(x))g(x) \xrightarrow[m \to \infty]{} 0,$$
$$|(s_m(x) - f(x))g(x)| \leqslant 2k|f(x)|.$$

因为 $\int_\Omega |f|dx \leqslant \|f\|_{L^p}(\mu(\Omega))^{\frac{1}{q}} < \infty$, 函数 f 属于空间 $L^1(\Omega)$. 将 Lebesgue 控制收敛定理 (定理 1.15-3) 应用于序列 $((s_m - f)g)_{m=1}^\infty$, 即得当 $m \to \infty$ 时, $\int_\Omega |(s_m(x) - f(x))g(x)|dx \to 0$. 这样, 一方面由 $|\ell(s_m) - \int_\Omega fgdx| \leqslant \int_\Omega |(s_m - f)g|dx$ 可得当 $m \to \infty$ 时

$$\ell(s_m) \to \int_\Omega fgdx.$$

另一方面, 对几乎所有的 $x \in \Omega$,

$$|s_m(x) - f(x)|^p \xrightarrow[m \to \infty]{} 0$$
$$|s_m(x) - f(x)|^p \leqslant 2^p|f(x)|^p,$$

而函数 $|f|^p$ 属于 $L^1(\Omega)$. 再次应用 Lebesgue 控制收敛定理, 即得当 $m \to \infty$ 时, $\|s_m - f\|_{L^p}^p \to 0$, 从而由 ℓ 的连续性, 当 $m \to \infty$ 时

$$\ell(s_m) \to \ell(f).$$

因此, $\ell(f) = \int_\Omega fgdx$.

(iv) 设 $\mu(\Omega) < \infty, p = 1$, 则在 (ii) 中出现的函数 $g \in L^1(\Omega)$ 满足

$$g \in L^\infty(\Omega) \text{ 且 } \|g\|_{L^\infty} \leqslant \|\ell\|_{(L^1)'}.$$

为简明起见, 记 $\|\ell\| := \|\ell\|_{(L^1)'}$. 给定 $\varepsilon > 0$, 定义集合

$$A_\varepsilon := \{x \in \Omega;\ |g(x)| \geqslant \|\ell\| + \varepsilon\}.$$

对给定的任意整数 $k \geqslant 1$, 定义函数 $f_k^\varepsilon \in L^1(\Omega)$ 为

$$f_k^\varepsilon(x) := \begin{cases} \operatorname{sgn} g(x), & x \in A_\varepsilon \cap B_k, \\ 0, & x \in \Omega - (A_\varepsilon \cap B_k), \end{cases}$$

这里集合 B_k 的定义同 (iii). 因为对所有的 $x \in A_\varepsilon \cap B_k$, $|g(x)| \geqslant \|\ell\| + \varepsilon$, 由 f_k^ε 的定义可知

$$\mu(A_\varepsilon \cap B_k)(\|\ell\| + \varepsilon) \leqslant \int_{A_\varepsilon \cap B_k} |g|dx = \int_\Omega f_k^\varepsilon g dx.$$

同时, 由 (iii) 又有

$$\int_\Omega f_k^\varepsilon g dx = \ell(f_k^\varepsilon) \leqslant \|\ell\| \, \|f_k^\varepsilon\|_{L^1(\Omega)} = \|\ell\|\mu(A_\varepsilon \cap B_k).$$

联合上述两个不等式, 即得对任何 $k \geqslant 1$,

$$\mu(A_\varepsilon \cap B_k) = 0.$$

因为函数 g 几乎处处有限 (注意到 $g \in L^1(\Omega)$), 所以集合 Ω 可以表为 $\Omega = (\cup_{k=1}^\infty B_k) \cup B$, 其中 $\mu(B) = 0$. 由关系式 $A_\varepsilon = (\cup_{k=1}^\infty (A_\varepsilon \cap B_k)) \cup (A \cap B)$, 所以 $\mu(A_\varepsilon) = 0$. 由 $\varepsilon > 0$ 的任意性, 得到 $\mu(\{x \in \Omega; |g(x)| \geqslant \|\ell\|\}) = \mu\left(\cup_{m=1}^\infty A_{\frac{1}{m}}\right) = 0$, 因此, 对几乎所有的 $x \in \Omega$, $|g(x)| \leqslant \|\ell\|$.

(v) 设 $\mu(\Omega) < \infty, 1 < p < \infty$, 则在 (ii) 中出现的函数 g 满足

$$g \in L^q(\Omega) \text{ 且 } \|g\|_{L^q} \leqslant \|\ell\|_{(L^p)'}.$$

定义集合 $A := \{x \in \Omega; g(x) \neq 0\}$, 给定任意整数 $k \geqslant 1$, 定义函数 $f_k : \Omega \to \mathbb{R}$ 为

$$f_k(x) := \begin{cases} \operatorname{sgn} g(x)|g(x)|^{q-1}, & x \in A \cap B_k, \\ 0, & x \in \Omega - (A \cap B_k), \end{cases}$$

这里 B_k 如 (iii) 中定义. 则由 f_k 的定义可知

$$\int_\Omega |f_k|^p dx = \int_{A \cap B_k} |g|^q dx \leqslant \int_{B_k} |g|^q dx;$$

从而由 (iii) 导出

$$\begin{aligned} \int_{B_k} |g^q| dx = \int_{B_k} f_k g dx = \int_\Omega f_k g dx \\ = \ell(f_k) \leqslant \|\ell\|_{(L^p)'} \|f_k\|_{L^p} \leqslant \|\ell\|_{(L^p)'} \left(\int_{B_k} |g|^q dx\right)^{\frac{1}{p}}. \end{aligned}$$

因此,

$$\left(\int_{B_k} |g|^q dx\right)^{1-\frac{1}{p}} = \left(\int_{B_k} |g|^q dx\right)^{\frac{1}{q}} \leqslant \|\ell\|_{(L^p)'}.$$

因为这个不等式对任何整数 $k \geqslant 1$ 均成立, 而 $\Omega = (\cup_{k=1}^\infty B_k) \cup B$, 其中 $\mu(B) = 0$, 令 $k \to \infty$ 即得

$$\|g\|_{L^q} = \lim_{k \to \infty} \left(\int_{B_k} |g|^q dx\right)^{\frac{1}{q}} \leqslant \|\ell\|_{(L^p)'}.$$

(vi) 设 $\mu(\Omega) < \infty, 1 \leqslant p < \infty$. 则对任意给定的 $\ell \in (L^p(\Omega))'$, 存在函数 $g \in L^q(\Omega)$, 使得对任意的 $f \in L^p(\Omega)$,

$$\ell(f) = \int_\Omega f g dx \text{ 且 } \|g\|_{L^q} = \|\ell\|_{(L^p)'}.$$

由 (ii), (iv) 和 (v) 可知, 对任何 $1 \leqslant p < \infty$, 存在函数 $g \in L^q(\Omega)$, 使得对任何可测的简单函数 s, $\ell(s) = \int_\Omega sgdx$.

给定函数 $f \in L^p(\Omega)$, 由定理 1.14-5, 存在可测的简单函数 $s_m : \Omega \to \mathbb{R}, m \geqslant 1$, 使得对所有的 $m \geqslant 1$, $|s_m| \leqslant |f|$, 且对几乎所有 $x \in \Omega$, $f(x) = \lim\limits_{m\to\infty} s_m(x)$. 将 Lebesgue 控制收敛定理应用于函数 $|s_m - f|^p$, $m \geqslant 1$, 因为它在 Ω 上几乎处处收敛于 0, 且以函数 $2^p|f|^p \in L^1(\Omega)$ 界于上, 从而当 $m \to \infty$ 时, $\|s_m - f\|_{L^p(\Omega)} \to 0$. 由 ℓ 的连续性, 当 $m \to \infty$ 时

$$\ell(s_m) \to \ell(f).$$

另一方面,

$$\begin{aligned}\left|\ell(s_m) - \int_\Omega fgdx\right| &= \left|\int_\Omega (s_m - f)gdx\right| \\ &\leqslant \|s_m - f\|_{L^p(\Omega)}\|g\|_{L^q(\Omega)}.\end{aligned}$$

因此, 当 $m \to \infty$ 时

$$\ell(s_m) \to \int_\Omega fgdx.$$

所以, 对所有的 $f \in L^p(\Omega), \ell(f) = \int_\Omega fgdx$.

由在 (i) 中得到的不等式 $\|\ell\|_{(L^p)'} \leqslant \|g\|_{L^q}$ 和在 (iv) (v) 中建立的不等式 $\|g\|_{L^q} \leqslant \|\ell\|_{(L^p)'}$, 即得 $\|g\|_{L^q} = \|\ell\|_{(L^p)'}$.

(vii) 最后, 设 $\mu(\Omega) = \infty$.

对每个整数 m, 令

$$\Omega_m := \left\{x \in \Omega;\ \text{dist}\,(x, \mathbb{R}^n - \Omega) > \frac{1}{m}\right\} \cap B(0, m).$$

因此,

$$\Omega = \bigcup_{m=1}^{\infty} \Omega_m, \quad \mu(\Omega_m) < \infty,$$
$$\Omega_m \subset \Omega_{m+1}, \quad m \geqslant 1.$$

任意给定函数 $f \in L^p(\Omega_m)$, 定义函数 $f^\# : \Omega \to [-\infty, \infty]$ 为 $f^\#|_{\Omega_m} := f, f^\#|_{\Omega - \Omega_m} := 0$, 则 $f^\#$ 属于 $L^p(\Omega)$. 这样, 对每个 $m \geqslant 1$, 令 $f \in L^p(\Omega)$

$$\ell_m(f) := \ell(f^\#),$$

就定义了 $L^p(\Omega_m)$ 上的一个连续线性泛函, 满足 $\|\ell_m\|_{(L^p(\Omega_m))'} \leqslant \|\ell\|_{(L^p)'}$. 另外, 因为 $\mu(\Omega_m) < \infty$, 由 (vi) 的结果可知, 存在函数 $g_m \in L^q(\Omega_m)$, 使得对所有的 $f \in L^p(\Omega_m)$,

$$\ell_m(f) = \int_{\Omega_m} g_m f dx, \text{ 且 } \|g_m\|_{L^q} = \|\ell_m\|_{(L^p(\Omega_m))'}.$$

任意给定函数 $f \in L^p(\Omega)$, 定义函数 $\tilde{f}: \Omega_{m+1} \to [-\infty, \infty]$ 为 $\tilde{f}|_{\Omega_m} := f, \tilde{f}|_{\Omega_{m+1}-\Omega_m} := 0$, 它满足 $\tilde{f}^\# = f^\#, \tilde{f}^\#|_{\Omega_{m+1}} = \tilde{f}, \tilde{f}^\#|_{\Omega-\Omega_{m+1}} = 0$. 因此, 对所有的 $f \in L^p(\Omega_m)$,

$$\begin{aligned}\int_{\Omega_m} g_m f dx &= \ell_m(f) = \ell(f^\#) = \ell(\tilde{f}^\#) \\ &= \ell_{m+1}(\tilde{f}) = \int_{\Omega_{m+1}} g_{m+1} \tilde{f} dx = \int_{\Omega_m} g_{m+1} f dx.\end{aligned}$$

另一方面, 函数 g_m 和 $g_{m+1}|\Omega_m$ 均属于空间 $L^q(\Omega_m)$. 因此, 对所有的 $f \in L^p(\Omega_m)$,

$$\int_{\Omega_m} (g_{m+1} - g_m) f dx = 0,$$

这个式子当然对一切 $\varphi \in \mathcal{D}(\Omega_m)$ 成立, 从而

$$g_{m+1} - g_m = 0$$

在 Ω 上几乎处处成立.

这样, 我们可以一意地定义函数 $g : \Omega \to [-\infty, \infty]$ 为对每个 $x \in \Omega, g(x) := g_{m(x)}(x)$, 其中 $m(x) := \min\{m \geqslant 1; x \in \Omega_m\}$. 这个函数显然满足对每个 $m \geqslant 1$,

$$g|_{\Omega_m} = g_m \in L^q(\Omega_m).$$

现在我们证明 $g \in L^q(\Omega)$. 如果 $p = 1$, 则 $q = \infty$, 显然有

$$\|q\|_{L^\infty} = \lim_{m\to\infty} \|g_m\|_{L^\infty} = \lim_{m\to\infty} \|\ell_m\|_{(L^1(\Omega_m))'} \leqslant \|\ell\|_{(L^1)'}.$$

如果 $1 < p < \infty$, 考察函数 $|g|^q \chi_{\Omega_m}, m \geqslant 1$. 对几乎所有的 $x \in \Omega$, 当 $m \to \infty$ 时

$$|g(x)|^q \chi_{\Omega_m}(x) \to |g(x)|^q.$$

对所有的 $m \geqslant 1$,

$$\begin{aligned}\int_\Omega |g|^q \chi_{\Omega_m} dx = \int_{\Omega_m} |g|^q dx &= \int_{\Omega_m} |g_m|^q dx \\ &= (\|\ell_m\|_{(L^p(\Omega_m))'})^q \leqslant (\|\ell\|_{(L^p)'})^q.\end{aligned}$$

对序列 $(|g|^q \chi_{\Omega_n})_{m\geqslant 1}$ 用 Fatou 引理 (定理 1.15-2), 即得

$$\int_\Omega |g|^q dx \leqslant \liminf_{m\to\infty} \int_\Omega |g_m|^q \chi_{\Omega_m} dx \leqslant (\|\ell\|_{(L^p)'})^q.$$

所以, 对一切 $1 \leqslant p < \infty$, 均有 $g \in L^q(\Omega)$, 且 $\|g\|_{L^q} \leqslant \|\ell\|_{(L^p)'}$.

最后, 给定函数 $f \in L^p(\Omega)$, 对每个 $m \geqslant 1$, 有

$$\begin{aligned}\ell(f\chi_{\Omega_m}) &= \ell_m(f|_{\Omega_m}) \\ &= \int_{\Omega_m} f|_{\Omega_m} g_m dx = \int_\Omega f\chi_{\Omega_m} g dx.\end{aligned}$$

因为当 $m \to \infty$ 时 $\|f\chi_{\Omega_m} - f\|_{L^p(\Omega)} \to 0$(利用 Lebesgue 控制收敛定理即可), 因此, 一方面当 $m \to \infty$ 时

$$\ell(f\chi_{\Omega_m}) \to \ell(f),$$

另一方面, 又因 $g \in L^q(\Omega)$, 故而当 $m \to \infty$ 时,

$$\int_\Omega (f\chi_{\Omega_m}) g dx \to \int_\Omega fg dx.$$

因此, 对任何 $f \in L^p(\Omega), \ell(f) = \int_\Omega fg dx$.

由在 (i) 中建立的不等式 $\|\ell\|_{(L^p)'} \leqslant \|g\|_{L^q}$ 和在上面建立的不等式 $\|g\|_{L^q} \leqslant \|\ell\|_{(L^p)'}$, 即得 $\|g\|_{L^q} = \|\ell\|_{(L^p)'}$. □

当 $p = 2$ 时, 定理 3.5-1 和 3.5-3 成为对任何 Hilbert 空间 (因而特别地对空间 ℓ^2 和 $L^2(\Omega)$) 都成立的一个一般结果的特殊情形. 这个一般结果也被称作 Hilbert 空间中的 F. Riesz 表示定理 (定理 4.6-1).

注　对于 $1 < p < \infty$, 基于 $L^p(\Omega)$ 的自反性 (自反空间定义于 5.14 节), 还可以给出定理 3.5-3 的一个特别漂亮的证明[5)].

最后, 考察 $p = \infty$ 的情况.

定理 3.5-4　*任意给定函数 $g \in L^1(\Omega)$, 由对所有的 $f \in L^\infty(\Omega)$, 关系式*

$$\ell(f) = \int_\Omega f(x)g(x)dx$$

定义了 $L^\infty(\Omega)$ 上的一个连续线性泛函, 而且

$$\|\ell\|_{(L^\infty(\Omega))'} = \|g\|_{L^1(\Omega)}.$$

这样定义的等距算子 $g \in L^1(\Omega) \to \ell \in (L^\infty(\Omega))'$ 是一个线性单射, 但不是满射, 即作为赋范向量空间, $L^1(\Omega)$ 只能恒同于空间 $(L^\infty(\Omega))'$ 的一个真子空间[6)].

证明　这里采用在定理 3.5-3 证明中同样简化的记号. 因为对任何 $f \in L^\infty(\Omega)$ 和 $g \in L^1(\Omega)$,

$$\left|\int_\Omega fg dx\right| \leqslant \|f\|_{L^\infty}\|g\|_{L^1}.$$

对任何 $f \in L^\infty(\Omega)$, 由 $\ell(f) = \int_\Omega fg dx$ 在 $L^\infty(\Omega)$ 上定义了一个连续性泛函 ℓ, 其范数满足 $\|\ell\|_{(L^\infty)'} \leqslant \|g\|_{L^1}$.

为建立相反方向的不等式, 设 $f \in L^\infty(\Omega)$ 定义为当 $x \in \Omega$ 时 $f(x) = \mathrm{sgn}\, g(x)$, 故 $\|f\|_{L^\infty} \leqslant 1$. 于是

$$\begin{aligned}\|g\|_{L^1} &= \int_\Omega |g|dx = \int_\Omega fg dx \\ &= \ell(f) \leqslant \|\ell\|_{(L^\infty)'}\|f\|_{L^\infty} \leqslant \|\ell\|_{(L^\infty)'}.\end{aligned}$$

[5)] 见 Brezis [2011, 定理 4.11].

[6)] 空间 $(L^\infty(\Omega))'$ 的完整描述在 Yosida [1965, 第 4 章第 9 节] 中给出.

因此, $\|\ell\|_{(L^\infty)'} = \|g\|_{L^1(\Omega)}$. 所以, 这样定义的显然是线性的映射 $g \in L^1(\Omega) \to \ell \in (L^\infty(\Omega))'$ 是等距算子 (从而是单射).

证明这个等距算子不是满射的最便捷的途径是注意到空间 $L^\infty(\Omega)$ 不是可分的 (定理 2.5-4), 余下采用定理 3.5-2 证明末尾部分的同样推理即可. □

3.6 Banach 空间的级数

现在, 我们把注意力放在 Banach 空间的另一值得注意的特点, 即这些空间中级数收敛的一个十分简单的充分条件. 但首先我们需要一些定义.

设 $(X, \|\cdot\|)$ 为赋范向量空间, $(x_n)_{n=1}^\infty$ 为向量 $x_n \in X$ 的序列. 称记号 $\sum_{n=1}^\infty x_n$ 为一个级数. 对每个整数 $k \geqslant 1$, 称

$$s_k := \sum_{n=1}^k x_n$$

为级数 $\sum_{n=1}^\infty x_n$ 的前 k 项的部分和. 如果序列 $(s_k)_{k=1}^\infty$ 在 X 中收敛, 则称级数 $\sum_{n=1}^\infty x_n$ 为收敛的. 此时, 记

$$\sum_{n=1}^\infty x_n = s,$$

其中 $s := \lim_{k\to\infty} s_k$, 称 s 为级数的和. 注意, 当级数收敛时, 记号 $\sum_{n=1}^\infty x_n$ 同时表示级数自身及其和.

关于级数收敛的下述的充分条件是基本的.

定理 3.6-1 (Banach 空间中级数的收敛性) 设 $(X, \|\cdot\|)$ 为 Banach 空间, $\sum_{n=1}^\infty x_n$ 为向量 $x_n \in X$ 的一个级数, 满足[7)]

$$\sum_{n=1}^\infty \|x_n\| < \infty.$$

则级数 $\sum_{n=1}^\infty x_n$ 收敛, 其和满足

$$\left\|\sum_{n=1}^\infty x_n\right\| \leqslant \sum_{n=1}^\infty \|x_n\|.$$

证明 因为级数 $\sum_{n=1}^\infty \|x_n\|$ 收敛, 记 $\sigma_k := \sum_{n=1}^k \|x_n\|$, 序列 $(\sigma_k)_{k=1}^\infty$ 是实数的 Cauchy 序列. 因为部分和 $s_k = \sum_{n=1}^k x_n, k \geqslant 1$, 满足对任何 $k \geqslant \ell + 1$,

$$\|s_k - s_l\| = \left\|\sum_{n=\ell+1}^k x_n\right\| \leqslant \sum_{n=\ell+1}^k \|x_n\| = \sigma_k - \sigma_\ell,$$

[7)] 假设读者熟悉实或复的数项级数的基本性质. 对给定的数 $\alpha_n \geqslant 0, n \geqslant 1$, 记号 $\sum_{n=1}^\infty \alpha_n < \infty$ 表示级数 $\sum_{n=1}^\infty \alpha_n$ 在 $\mathbb{R}$ 中收敛.

故而序列 $(s_k)_{k=1}^{\infty}$ 是 Banach 空间 $(X, \|\cdot\|)$ 中的 Cauchy 序列, 所以在 X 中收敛. 同时, 极限 s 满足

$$\|s\| = \lim_{k\to\infty} \|s_k\|.$$

$$\|s_k\| \leqslant \sum_{n=1}^{k} \|x_n\| \leqslant \sum_{n=1}^{\infty} \|x_n\|.$$ □

这个结果的第一个应用是给出一个简单的充分条件, 以利用 Neumann 级数[8] $\sum_{n=0}^{\infty} A^n$ (其中 $A^0 = I$) 定义一个具有特殊形式的线性算子 $(I - A)$ 的逆. 下面定理把熟知的公式

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n,$$

其中 $|z| < 1, z \in \mathbb{C}$, 推广到一般的 Banach 空间.

定理 3.6-2 (Neumann 级数的收敛) 设 $(X, \|\cdot\|)$ 为 Banach 空间, $A \in \mathcal{L}(X)$ 满足

$$\|A\| < 1,$$

其中 $\|A\|$ 表示 A 的算子范数 (2.9 节) 则连续线性算子 $(I - A) : X \to X$ 是双射, 其逆 $(I - A)^{-1} : X \to X$ 也是连续线性算子. 此外,

$$(I - A)^{-1} = \sum_{n=0}^{\infty} A^n \text{ 且 } \|(I - A)^{-1}\| \leqslant \frac{1}{1 - \|A\|}.$$

证明 由不等式 $\|A\| < 1$ 和对每个 $n \geqslant 0$, $\|A^n\| \leqslant \|A\|^n$ (定理 2.9-5(d)) 可得

$$\sum_{n=0}^{\infty} \|A^n\| \leqslant \sum_{n=0}^{\infty} \|A\|^n < \infty.$$

因为 $\mathcal{L}(X)$ 为 Banach 空间 (定理 3.2-4), 由定理 3.6-1, 级数 $\sum_{n=0}^{\infty} A^n$ 在 $\mathcal{L}(X)$ 中收敛. 用 $B \in \mathcal{L}(X)$ 记其和, 即

$$B = \sum_{n=0}^{\infty} A^n := \lim_{k\to\infty} B_k, \text{ 其中 } B_k := \sum_{n=0}^{k} A^n.$$

则

$$AB = \lim_{k\to\infty} AB_k = \lim_{k\to\infty} (B_{k+1} - I) = B - I,$$
$$BA = \lim_{k\to\infty} B_k A = \lim_{k\to\infty} (B_{k+1} - I) = B - I,$$

所以

$$I = B(I - A) = (I - A)B.$$

[8] 这个命名为纪念 Carl Neumann (1832—1925).

因此, $(I-A) \in \mathcal{L}(X)$ 是双射 (因为 $(I-A)$ 有左逆和右逆), 且

$$(I-A)^{-1} = B = \sum_{n=0}^{\infty} A^n.$$

此外, 由定理 3.6-1,

$$\|(I-A)^{-1}\| \leqslant \sum_{n=0}^{\infty} \|A^n\| \leqslant \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1-\|A\|}. \qquad \square$$

作为定理 3.6-2 的第一个应用, 我们来建立从 Banach 空间到赋范向量空间上具有连续逆的连续线性算子的一个重要性质.

定理 3.6-3 设 X 为 Banach 空间, Y 为赋范向量空间, 则集合

$$\mathcal{U} := \{A \in \mathcal{L}(X;Y); A: X \to Y \text{ 为双射, 且 } A^{-1} \in \mathcal{L}(Y;X)\}$$

是赋范向量空间 $(\mathcal{L}(X;Y), \|\cdot\|_{\mathcal{L}(X;Y)})$ 中的开集.

又设 $A \in \mathcal{U}$, 如果

$$\|B-A\| < \frac{1}{\|A^{-1}\|},$$

则 $B \in \mathcal{U}$, 此时

$$\|B^{-1}\| \leqslant \frac{\|A^{-1}\|}{1-\|A^{-1}(B-A)\|} \leqslant \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|B-A\|},$$

$$\|B^{-1}-A^{-1}\| \leqslant \frac{\|A^{-1}\|^2\|B-A\|}{1-\|A^{-1}(B-A)\|} \leqslant \frac{\|A^{-1}\|^2\|B-A\|}{1-\|A^{-1}\|\|B-A\|},$$

因而, 映射 $A \in \mathcal{U} \to A^{-1} \in \mathcal{U}$ 是连续的.

证明 设 $A \in \mathcal{U}$. 因为 $\mathcal{L}(X)$ 是 Banach 空间 (此因 X 为 Banach 空间, 参见定理 3.2-4). 如果

$$\|B-A\|_{\mathcal{L}(X;Y)} < \frac{1}{\|A^{-1}\|_{\mathcal{L}(Y;X)}},$$

于是 $\|A^{-1}(B-A)\|_{\mathcal{L}(X)} < 1$, 从而由定理 3.6-2 可知 $(I_X + A^{-1}(B-A)) \in \mathcal{L}(X)$ 是具有连续逆的双射. 因此, 当 $\|B-A\| < (\|A^{-1}\|)^{-1}$ 时,

$$B = A(I + A^{-1}(B-A)) \in \mathcal{L}(X;Y)$$

也是具有连续逆的双射, 且其逆为

$$B^{-1} = (I_X + A^{-1}(B-A))^{-1}A^{-1} \in \mathcal{L}(Y;X),$$

所以 $\mathcal{U}$ 是 $\mathcal{L}(X;Y)$ 中的开集. 另外, 由 B^{-1} 的上述表达式和定理 3.6-2 可知, 当 $\|B-A\| < \frac{1}{\|A^{-1}\|}$ 时

$$\|B^{-1}\| \leqslant \frac{\|A^{-1}\|}{1-\|A^{-1}(B-A)\|} \leqslant \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|B-A\|}.$$

又当 $\|B-A\| < \frac{1}{\|A^{-1}\|}$ 时, 由等式 $B^{-1}-A^{-1}=B^{-1}(A-B)A^{-1}$ 可得

$$\|B^{-1}-A^{-1}\| \leqslant \frac{\|A^{-1}\|^2\|B-A\|}{1-\|A^{-1}\|\|B-A\|}.$$ □

注 (1) 后面将会证明映射 $A \in \mathcal{U} \subset \mathcal{L}(X;Y) \to A^{-1} \in \mathcal{L}(Y;X)$ 不仅是连续的, 而且实际上是无限阶可微的 (定理 7.12-2).

(2) 如果集合 $\mathcal{U}$ 非空, 则空间 Y 必定是 Banach 空间.

在赋范向量空间 $(X, \|\cdot\|)$ 中, 如果 $x_n \in X, n \geqslant 1$ 满足

$$\sum_{n=1}^{\infty} \|x_n\| < \infty,$$

则称级数 $\sum_{n=1}^{\infty} x_n$ 为绝对收敛的. 由定理 3.6-1 可知, Banach 空间中任何绝对收敛的级数都是收敛的. 值得注意的是具有这个性质的空间一定是 Banach 空间, 这就给出了制定一个赋范向量空间是 Banach 空间的准则 (下面的定理 3.6-5 说明了这一点).

定理 3.6-4 设 $(X, \|\cdot\|)$ 为赋范向量空间, 这个空间中每个绝对收敛级数均收敛, 则 X 为 Banach 空间.

证明 设 $(x_n)_{n=1}^{\infty}$ 为 X 中的 Cauchy 序列, 所以存在子列 $(x_{\sigma(n)})_{n=1}^{\infty}$, 使得当 $n \geqslant 1$ 时 $\|x_{\sigma(n+1)} - x_{\sigma(n)}\| \leqslant \frac{1}{2^n}$, 从而 $\sum_{n=1}^{\infty} \|x_{\sigma(n+1)} - x_{\sigma(n)}\| < \infty$. 由假设, 存在元素 $x \in X$ 使得

$$x = \lim_{k\to\infty} \sum_{n=1}^{k} (x_{\sigma(n+1)} - x_{\sigma(n)}) = \lim_{k\to\infty} (x_{\sigma(k+1)} - x_{\sigma(1)}).$$

因此, 子列 $(x_{\sigma(n)})_{n=1}^{\infty}$ 收敛. 但是, 包含收敛子列的 Cauchy 序列是收敛的 (定理 1.12-1(c)). □

将上述判别准则应用于商空间 (2.2 节), 可得

定理 3.6-5 设 X 为 Banach 空间, Z 为 X 的闭子空间. 则商空间 (其上赋以商范数, 见定理 2.2-3) 也是 Banach 空间.

证明 设 $\sum_{n=1}^{\infty}[x_n]$ 是商空间 X/Z 中绝对收敛的级数, 由商范数的定义 (定理 2.2-3), 对任意的 $n \geqslant 1$, 存在 $y_n \in [x_n]$, 使得 $\|y_n\| \leqslant \|[x_n]\| + \frac{1}{2^n}$. 因此, $\sum_{n=1}^{\infty}\|y_n\| \leqslant 1 + \sum_{n=1}^{\infty}\|[x_n]\| < \infty$. 因为 X 是完备的, 所以存在 $x \in X$, 使得 $x = \lim_{k\to\infty}\sum_{n=1}^{k} y_n$. 现在

$$\left[x - \sum_{n=1}^{k} y_n\right] = [x] - \sum_{n=1}^{k}[y_n] = [x] - \sum_{n=1}^{k}[x_n],$$

从而

$$\left\|[x] - \sum_{n=1}^{k}[x_n]\right\| = \left\|\left[x - \sum_{n=1}^{k} y_n\right]\right\| \leqslant \left\|x - \sum_{n=1}^{k} y_n\right\|.$$

由此可知级数 $\sum_{n=1}^{\infty}[x_n]$ 收敛于 $[x]$. 于是, 由定理 3.6-4, X/Z 为 Banach 空间. □

习题

3.6-1 设 $\boldsymbol{A}$ 为 $N\times N$ 实矩阵.

(1) 证明级数 $\sum_{n=0}^{\infty}\frac{1}{n!}\boldsymbol{A}^n$ 在由所有 $N\times N$ 实矩阵组成的向量空间 $\mathbb{M}^N$ 中收敛. 记其和为 $e^{\boldsymbol{A}}:=\sum_{n=0}^{\infty}\frac{1}{n!}\boldsymbol{A}^n$, 称作指数为 $\boldsymbol{A}$ 的矩阵.

(2) 证明

$$e^{\boldsymbol{A}}=\lim_{k\to\infty}\left(\boldsymbol{I}+\frac{\boldsymbol{A}}{k}\right)^k.$$

(3) 证明 $\det(e^{\boldsymbol{A}})=e^{\operatorname{tr}\boldsymbol{A}}$ (特别地, 由此说明矩阵 $e^{\boldsymbol{A}}$ 总是可逆的).

(4) 设 $\boldsymbol{B}$ 为实 $N\times N$ 矩阵, 证明: 如果 $\boldsymbol{A}$ 和 $\boldsymbol{B}$ 可交换, 则 $e^{(\boldsymbol{A}+\boldsymbol{B})}=e^{\boldsymbol{A}}e^{\boldsymbol{B}}$, 特别地, 可得 $e^{\boldsymbol{A}}$ 和 $e^{\boldsymbol{B}}$ 也可交换.

3.6-2 (1) 给定实 $N\times N$ 矩阵 $\boldsymbol{A}$ 和向量 $\boldsymbol{u}_0\in\mathbb{R}^n$, 对 $t\geqslant 0$, 令 $\boldsymbol{u}(t)=(u_i(t))_{i=1}^N:=e^{t\boldsymbol{A}}\boldsymbol{u}_0\in\mathbb{R}^N$, 其中 $e^{t\boldsymbol{A}}$ 表示指数为 $t\boldsymbol{A}$ 的矩阵 (习题 3.6-1). 证明: 对每个 $1\leqslant i\leqslant N$, 函数 $t\in[0,\infty[\to u_i(t)$ 均为可微的, 且

$$\boldsymbol{u}'(t)=\boldsymbol{A}\boldsymbol{u}(t),t\geqslant 0;\boldsymbol{u}(0)=\boldsymbol{u}_0,$$

其中 $\boldsymbol{u}'(t):=(u_i'(t))_{i=1}^N$.

(2) 对给定的向量场 $\boldsymbol{b}\in\mathcal{C}([0,\infty[;\mathbb{R}^N)$, 求

$$\boldsymbol{u}'(t)=\boldsymbol{A}\boldsymbol{u}(t)+\boldsymbol{b}(t),t\geqslant 0,\boldsymbol{u}(0)=\boldsymbol{0}$$

的解 $t\in[0,\infty[\to\boldsymbol{u}(t)=(u_i(t))_{i=1}^N$ 的显式表示.

这样, 问题 (1) 和 (2) 对于特定的线性常微分方程的 Cauchy 问题, 提供了显式解. 我们将在 3.8 和 3.11 节中建立非线性常微分方程 Cauchy 问题解的存在性.

3.6-3 设 $(X,\|\cdot\|)$ 为 Banach 空间, $A\in\mathcal{L}(X)$ 满足对某个 $p\geqslant 2$, 有 $\|A^p\|<1$. 证明 $(I-A)\in\mathcal{L}(X)$ 为双射, 且其逆 $(I-A)^{-1}:X\to X$ 也是连续的.

3.7 Banach 不动点定理

设 $f:X\to X$ 是集合 X 到其自身的一个映射. 所谓 f 的不动点是指满足

$$f(x)=x$$

的任意点 $x\in X$.

设 (X,d) 为一距离空间. 对映射 $f:X\to X$, 如果存在常数 k, 使得 $0<k<1$, 且对任何 $x,y\in X$, 均有

$$d(f(x),f(y))\leqslant kd(x,y),$$

则称 f 为压缩映射.

下面这个由 Stefan Banach 得到的定理[9] 是分析中最重要的结果之一. 其证明是简单的, 但却有许多关键的应用, 如线性方程求解时迭代方法的收敛性 (习题 3.7-6), 常微分方程解的存在性 (定理 3.8-1), 两点边值问题 (定理 3.9-1), Lax-Milgram 引理 (定理 6.2-1) 或隐函数定理 (定理 7.12-1), 还可以举出很多.

尽管本章讨论的是 Banach 空间, 我们还是在完备距离空间中证明这个定理 (实际上在这个更一般的空间上的证明与在 Banach 空间上的证明是相同的). 还有一些有趣的补充将在习题 3.7-1—3.7-5 中给出.

定理 3.7-1 (Banach 不动点定理) 设 (X, d) 为完备的距离空间. 则任何压缩映射 $f : X \to X$ 有且仅有一个不动点 $x \in X$.

此外, 任意给定点 $x_0 \in X$, 由

$$x_{n+1} = f(x_n), \quad n \geqslant 0,$$

定义的序列 $(x_n)_{n=0}^{\infty}$ 当 $n \to \infty$ 时, 收敛于 x, 且成立以下估计:

$$d(x_n, x) \leqslant Ck^n, \quad n \geqslant 0, \text{ 其中 } C := \frac{d(f(x_0), x_0)}{1-k}.$$

证明 记 X 上的距离为 d. 对任意给定的 $x_0 \in X$, 由 $x_{n+1} = f(x_n), n \geqslant 0$ 定义的序列 $(x_n)_{n=0}^{\infty}$ 是一个 Cauchy 序列. 这是因为对任何 $p \geqslant 1$,

$$d(x_{p+1}, x_p) \leqslant kd(x_p, x_{p-1}) \leqslant \cdots \leqslant k^p d(x_1, x_0),$$

故对任意的 $m > n \geqslant 0$,

$$\begin{aligned} d(x_m, x_n) &\leqslant \sum_{p=n}^{m-1} d(x_{p+1}, x_p) \leqslant \left(\sum_{p=n}^{m-1} k^p\right) d(x_1, x_0). \\ &\leqslant k^n \left(\sum_{p=0}^{m-n-1} k^p\right) d(x_1, x_0) \leqslant \frac{k^n}{1-k} d(x_1, x_0). \end{aligned}$$

空间 (X, d) 是完备的, 所以存在 $x \in X$, 使得 $\lim_{n\to\infty} x_n = x$. 因为压缩映射显然是连续的, 故而

$$f(x) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x.$$

因此, x 是 f 的不动点. 设 $y \in X$ 也是 f 的一个不动点, 则

$$d(x, y) = d(f(x), f(y)) \leqslant kd(x, y).$$

因为 $k < 1$ 所以 $y = x$, 即 f 的不动点是唯一的. □

[9] S. Banach [1922]: Sur les opérations dans les ensembles abstraits et leurs applications aux équations intégrales. Fundamenta Mathematicae **3**, 133–181.

序列 $(x_n)_{n=0}^{\infty}$ 构造为, x_0 为 X 中的任意点, $x_{n+1}=f(x_n), n \geqslant 0$, 利用序列 $(x_n)_{n=0}^{\infty}$ 逼近 f 的不动点的方法称之为*逐次逼近法*或 *Picard 方法*[10].

如习题 3.7-6 将要说明的, Picard 方法实际上是解线性方程组的某些最基本的迭代方法的核心.

注 下面简单的反例将说明定理 3.7-1 中所有的假设都有本质意义:

压缩映射 $x \to \frac{x}{2}$ 在不完备的距离空间 $]0,\infty[$ 上没有不动点.

由完备距离空间 $X := [1,\infty[$ 到自身的映射 $f: x \to x+\frac{1}{x}$ 对任意的 $x,y \in X, x \neq y$ 满足 $d(f(x),f(y)) < d(x,y)$, 而 f 没有不动点. 但是要注意, 如果这样的映射定义于一个紧的距离空间上, 则它确有不动点, 参见习题 3.7-3.

在任何至少有两个点的距离空间 (完备或不完备) 中, X 上的恒等映射 f 满足对任何 $x,y \in X, d(f(x),f(y)) \leqslant d(x,y)$, 但它有超过一个的不动点.

习题

3.7-1 设 (X,d) 为完备距离空间, T 为拓扑空间, $(f_t)_{t\in T}$ 为一族满足下述性质的映射 $f_t: X \to X$: 对每个 $x \in X$, 映射 $t \in T \to f_t(x) \in X$ 是连续的, 而且存在常数 $k, 0<k<1$, 使得对任何 $x,y \in X$ 和 $t \in T$,

$$d(f_t(x),f_t(y)) \leqslant kd(x,y).$$

对每个 $t \in T$, 用 $x_t \in X$ 表示映射 f_t 的唯一的不动点. 证明映射 $t \in T \to x_t \in X$ 是连续的.

3.7-2 设 (X,d) 是完备距离空间, 映射 $f: X \to X$ 满足对某 $p \geqslant 2$, 含有 p 个因子的复合映射 $f \circ f \circ \cdots \circ f$ 为压缩映射. 注意, 这里并不假设映射 f 是连续的.

(1) 证明 f 有且仅有一个不动点 x.

(2) 证明: 任意给定 $x_0 \in X$, 由 $x_{n+1}=f(x_n), n \geqslant 0$, 定义的序列 $(x_n)_{n=0}^{\infty}$ 收敛于 x.

3.7-3 设 (X,d) 为紧距离空间 (因此 (X,d) 是完备的, 见定理 1.13-3), 映射 $f: X \to X$ 满足对任意的 $x,y \in X, x \neq y$ 有 $d(f(x),f(y)) < d(x,y)$.

(1) 证明 f 有且仅有一个不动点.

(2) 找一个这样的 f 未必是压缩映射的例子.

3.7-4 设 (X,d) 为紧距离空间, $f: X \to X$ 为连续映射, 对任意的 $x,y \in X$, 满足 $d(f(x),f(y)) \geqslant d(x,y)$. 证明: 对任意的 $x,y \in X$, 均有 $d(f(x),f(y)) = d(x,y)$.

3.7-5 设 $k>0$, A 是距离空间 (X,d) 的一个子集, 设映射 $f: A \to \mathbb{R}$ 满足 $|f(x)-f(y)| \leqslant kd(x,y)$ 对所有的 $x,y \in A$ 成立. 证明存在映射 $\tilde{f}: X \to \mathbb{R}$, 使得对所有的 $x \in A$, 有 $\tilde{f}(x)=f(x)$, 且对所有的 $x,y \in X$, 有 $|\tilde{f}(x)-\tilde{f}(y)| \leqslant kd(x,y)$. 这个结论即 MacShane 引

---

[10] 这个方法是在解两点边值问题 (习题 3.9-1 中将考察其形式) 时引入的, 见:

E. Picard [1893]: Sur l'application des méthodes d'approximations successives à l'étude de certaines équations différentielles ordinaires. Journal de Mathématiques Pures et Appliguées **9**, 217–271.

理[11].

3.7-6 (1) 考察形为 $\boldsymbol{u}=\boldsymbol{B}\boldsymbol{u}+\boldsymbol{c}$ 的线性系统, 这里 $\boldsymbol{B}$ 是 $N\times N$ 矩阵, $\boldsymbol{c}\in\mathbb{R}^N$, 且设对某矩阵范数 $\|\cdot\|$, $\|\boldsymbol{B}\|<1$. 证明这个线性系统有唯一解 $\boldsymbol{u}$, 定义

$$\boldsymbol{u}^{n+1}=\boldsymbol{B}\boldsymbol{u}^n+\boldsymbol{c},\quad n\geqslant 0,$$

其中 $\boldsymbol{u}^0\in\mathbb{R}^N$ 是任意向量, 证明向量 $\boldsymbol{u}^n\in\mathbb{R}^N$ 的序列 $(\boldsymbol{u}^n)_{n=0}^{\infty}$ 当 $n\to\infty$ 时收敛于 $\boldsymbol{u}$.

给定一个 $N\times N$ 矩阵 $\boldsymbol{A}=(a_{ij})$. 定义矩阵 $\boldsymbol{D},\boldsymbol{E}$ 和 $\boldsymbol{F}$ 分别为: $(\boldsymbol{D})_{ij}=a_{ij}\delta_{ij}$; 当 $i>j$ 时, $(-\boldsymbol{E})_{ij}=a_{ij}$, 当 $i\leqslant j$ 时, $(-\boldsymbol{E})_{ij}=0$; 当 $i<j$ 时 $(-\boldsymbol{F})_{ij}=a_{ij}$, 当 $i\geqslant j$ 时, $(-\boldsymbol{F})_{ij}=0$, 在上述足标中 $1\leqslant i,j\leqslant N$. 注意, $\boldsymbol{A}=\boldsymbol{D}-\boldsymbol{E}-\boldsymbol{F}$.

(2) 考察线性系统 $\boldsymbol{A}\boldsymbol{u}=\boldsymbol{b}$, 这里 $\boldsymbol{A}=(a_{ij})$ 为实 $N\times N$ 阵, 且 $a_{ii}\neq 0, 1\leqslant i\leqslant N, \boldsymbol{b}=(b_i)$ 是 $\mathbb{R}^N$ 中的向量. 所谓 Jacobi 方法是计算这种系统的解 $\boldsymbol{u}\in\mathbb{R}^N$ 的最简单的迭代方法; 任意给定向量 $\boldsymbol{u}^0\in\mathbb{R}^N$, 定义向量 $\boldsymbol{u}^n\in\mathbb{R}^N$ 的序列 $(\boldsymbol{u}^n)_{n=0}^{\infty}$ 为

$$\boldsymbol{D}\boldsymbol{u}^{n+1}=(\boldsymbol{E}+\boldsymbol{F})\boldsymbol{u}^n+\boldsymbol{b},\quad n\geqslant 0.$$

证明: 如果矩阵 $\boldsymbol{A}=(a_{ij})$ 是严格对角控制的, 即 $|a_{ii}|>\sum_{\substack{j=1\\ j\neq i}}^{N}|a_{ij}|, 1\leqslant i\leqslant N$, 则矩阵 $\boldsymbol{A}$ 可逆, 而且对任意给定的向量 $\boldsymbol{u}^0\in\mathbb{R}^N$, 当 $n\to\infty$ 时, 上面的序列 $(\boldsymbol{u}^n)_{n=0}^{\infty}$ 收敛于 $\boldsymbol{u}=\boldsymbol{A}^{-1}\boldsymbol{b}$.

提示: 利用矩阵范数 $\|\cdot\|_\infty$ (习题 2.9-1).

(3) 解线性方程组 $\boldsymbol{A}\boldsymbol{u}=\boldsymbol{b}$ 的 Gauss-Seidel 方法, 其中 $\boldsymbol{b}$ 是 $\mathbb{R}^N$ 中的向量: 定义向量 $\boldsymbol{u}^n\in\mathbb{R}^N$ 的序列 $(\boldsymbol{u}^n)_{n=0}^{\infty}$ 为

$$(\boldsymbol{D}-\boldsymbol{E})\boldsymbol{u}^{n+1}=\boldsymbol{F}\boldsymbol{u}^n+\boldsymbol{b},\quad n\geqslant 0,$$

其中 $\boldsymbol{u}^0$ 为 $\mathbb{R}^N$ 中任意向量. 证明: 如果矩阵 $\boldsymbol{A}$ 是严格对角控制的, 则 Gauss-Seidel 方法收敛.

提示: 证明矩阵 $(\boldsymbol{D}-\boldsymbol{E})^{-1}\boldsymbol{F}$ 的任何特征值 λ 均满足 $|\lambda|<1$, 再利用习题 2.9-3(1).

(4) 给定实 $N\times N$ 阵 $\boldsymbol{A}=(a_{ij})$, 其中 $a_{ii}\neq 0, 1\leqslant i\leqslant N$, 给定参数 $\omega\neq 0$. 所谓用松弛方法解线性方程组 $\boldsymbol{A}\boldsymbol{u}=\boldsymbol{b}$ 就是构造向量 $\boldsymbol{u}^n\in\mathbb{R}^N$ 的序列 $(\boldsymbol{u}^n)_{n=0}^{\infty}$ 为

$$\left(\frac{1}{\omega}\boldsymbol{D}-\boldsymbol{E}\right)\boldsymbol{u}^{n+1}=\left(\frac{1-\omega}{\omega}\boldsymbol{D}+\boldsymbol{F}\right)\boldsymbol{u}^n+\boldsymbol{b},\quad n\geqslant 0,$$

其中 $\boldsymbol{u}^0$ 是 $\mathbb{R}^N$ 中的任意向量 (Gauss-Seidel 方法相当于这里的 $\omega=1$ 的特殊情况). 证明: 如果 $0<\omega<2$, 矩阵 $\boldsymbol{A}$ 是对称且正定的, 则松弛方法中的序列 $(\boldsymbol{u}^n)_{n=0}^{\infty}$ 收敛. 这个结果组成了 Ostrowski-Reich[12] 定理.

提示: 证明 $\left\|\left(\frac{1}{\omega}\boldsymbol{D}-\boldsymbol{E}\right)^{-1}\left(\frac{1-\omega}{\omega}\boldsymbol{D}+\boldsymbol{F}\right)\right\|<1$, 这里 $\|\cdot\|$ 是从属于向量范数 $\boldsymbol{v}\in\mathbb{R}^N\to(\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{v})^{\frac{1}{2}}$ 的矩阵范数.

---

[11] E. J. MacShane [1934]: Extension of range of functions. Bulletin of the American Mathematical Society **40**, 837–842.

[12] A. M. Ostrowski [1954]: On the linear iteration procedures for symmetric matrices. Rendiconti Lincei-Matematicae Applicazioni **14**, 140–163.

E. Reich [1949]: On the convergence of the classical iterative method of solving linear simultaneous equatious. Annals of Mathematical Statistics **20**, 448–451.

注意, 当线性系统的矩阵是对角阵时, Jacobi 方法的一迭代包含了线性系统的解, 当线性系统的矩阵是下三角阵时, Gauss-Seidel 或松弛方法的一迭代包含了相应的解.[13)]

3.8 Banach 不动点定理的应用: 非线性常微分方程解的存在性; Cauchy-Lipschitz 定理; 单摆方程

设 K 为紧集, Y 为 Banach 空间. 作为 Banach 不动点定理在形为 $\mathcal{C}(K;Y)$ 的 Banach 空间 (定理 3.2-2) 上的第一个应用, 我们将对一类特殊的常微分方程组的初值问题或 Cauchy 问题, 建立解的存在性与唯一性[14)].

因为在应用中, 变量经常是时间, 所以用 t 表示. 不失一般性, 我们假设 "初始时刻" 是 $t=0$ (如果初始时刻 $t=t_0$, 则将 $(t-t_0)$ 作为新的 "时间变量" 即可). 所有连续可微映射 $\boldsymbol{v}:=(v_i)_{i=1}^N:[0,T]\to\mathbb{R}^N$ 组成的空间记作 $\mathcal{C}^1([0,T];\mathbb{R}^N)$, 对每个 $t\in[0,T]$, 记号 $\boldsymbol{v}'(t)$ 表示向量 $(v_i'(t))_{i=1}^N$.

定理 3.8-1 (Cauchy-Lipschitz 定理) 设 $\|\cdot\|$ 为 $\mathbb{R}^N$ 上的任意一个范数, $T>0, \boldsymbol{g}\in\mathcal{C}([0,T]\times\mathbb{R}^N;\mathbb{R}^N)$ 为一个映射, 且存在常数 $\gamma>0$, 使得

$$\|\boldsymbol{g}(t,\boldsymbol{w})-\boldsymbol{g}(t,\boldsymbol{v})\|\leqslant\gamma\|\boldsymbol{w}-\boldsymbol{v}\|$$

对一切 $t\in[0,T]$ 和一切 $\boldsymbol{w},\boldsymbol{v}\in\mathbb{R}^N$ 成立. 又设 $\boldsymbol{u}_0\in\mathbb{R}^N$ 为给定的向量. 则初始问题或 Cauchy 问题

$$\begin{aligned}\boldsymbol{u}'(t)&=\boldsymbol{g}(t,\boldsymbol{u}(t)),\quad 0\leqslant t\leqslant T,\\ \boldsymbol{u}(0)&=\boldsymbol{u}_0\end{aligned}$$

有且仅有一个解 $\boldsymbol{u}\in\mathcal{C}^1([0,T];\mathbb{R}^N)$.

证明 (i) 容易验证: 如果 $\boldsymbol{u}\in\mathcal{C}([0,T];\mathbb{R}^N)$ 是积分方程

$$\boldsymbol{u}(t)=\boldsymbol{u}_0+\int_0^t\boldsymbol{g}(s,\boldsymbol{u}(s))ds,\quad 0\leqslant t\leqslant T$$

的解, 则 $\boldsymbol{u}\in\mathcal{C}^1([0,T];\mathbb{R}^N)$, 且 $\boldsymbol{u}$ 是初值问题的解; 反之, 如果 $\boldsymbol{u}\in\mathcal{C}^1([0,T];\mathbb{R}^N)$ 是初值问题的解, 则 $\boldsymbol{u}$ 是积分方程的解.

(ii) 空间 $\mathcal{C}([0,T];\mathbb{R}^N)$ 上赋以范数

$$|||\cdot|||:\boldsymbol{v}\in\mathcal{C}([0,T];\mathbb{R}^N)\to\sup_{0\leqslant t\leqslant T}(e^{-\gamma t}\|\boldsymbol{v}(t)\|)$$

---

13) 用迭代方法解线性系统的一个特别有启发性的处理, 以及有关这里叙述的迭代方法的参考文献可见 VARGA [1962].

14) 关于常微分方程的进一步处理, 包括历史参考及文献目录可参见两本经典著作 CODDINGTON & LEVINSON [1955] 和 HARTMAN [2002].

后是一个 Banach 空间 (因为在此空间上这一范数显然等价于 sup 范数; 参见定理 3.2-2). 关于这个范数, 由

$$\boldsymbol{F}(\boldsymbol{v})(t)=\boldsymbol{u}_0+\int_0^t \boldsymbol{g}(s,\boldsymbol{v}(s))ds, \quad 0\leqslant t\leqslant T$$

定义的映射 $\boldsymbol{F}:\mathcal{C}([0,T];\mathbb{R}^N)\to\mathcal{C}([0,T];\mathbb{R}^N)$ 是一个压缩映射.

为说明这一点, 对一切 $\boldsymbol{v},\boldsymbol{w}\in\mathcal{C}([0,T];\mathbb{R}^N)$, 有当 $0\leqslant t\leqslant T$ 时

$$(\boldsymbol{F}(\boldsymbol{w})-\boldsymbol{F}(\boldsymbol{v}))(t)=\int_0^t e^{\gamma s}e^{-\gamma s}(\boldsymbol{g}(s,\boldsymbol{w}(s))-\boldsymbol{g}(s,\boldsymbol{v}(s)))ds.$$

由此可得当 $0\leqslant t\leqslant T$ 时

$$\begin{aligned}\|(\boldsymbol{F}(\boldsymbol{w})-\boldsymbol{F}(\boldsymbol{v}))(t)\| &\leqslant \left(\int_0^t e^{\gamma s}ds\right)\sup_{0\leqslant s\leqslant T}(e^{-\gamma s}\|\boldsymbol{g}(s,\boldsymbol{w}(s))-\boldsymbol{g}(s,\boldsymbol{v}(s))\|)\\ &\leqslant \gamma\left(\int_0^t e^{\gamma s}ds\right)|||\boldsymbol{w}-\boldsymbol{v}|||\\ &\leqslant e^{\gamma t}(1-e^{-\gamma T})|||\boldsymbol{w}-\boldsymbol{v}|||,\end{aligned}$$

这里利用了 $\gamma\int_0^t e^{\gamma s}ds=e^{\gamma t}-1=e^{\gamma t}(1-e^{-\gamma t})\leqslant e^{\gamma t}(1-e^{-\gamma T})$. 因此,

$$\begin{aligned}|||\boldsymbol{F}(\boldsymbol{w})-\boldsymbol{F}(\boldsymbol{v})||| &= \sup_{0\leqslant t\leqslant T}(e^{-\gamma t}\|(\boldsymbol{F}(\boldsymbol{w})-\boldsymbol{F}(\boldsymbol{v}))(t)\|)\\ &\leqslant (1-e^{-\gamma T})|||\boldsymbol{w}-\boldsymbol{v}|||.\end{aligned}$$

于是, $\boldsymbol{F}$ 是压缩映射. 由 Banach 不动点定理 (定理 3.7-1), 这个压缩映射有且仅有一个不动点 $\boldsymbol{u}$, 即函数 $\boldsymbol{u}\in\mathcal{C}([0,T];\mathbb{R}^N)$ 满足当 $0\leqslant t\leqslant T$ 时

$$\boldsymbol{u}(t)=\boldsymbol{u}_0+\int_0^t \boldsymbol{g}(s,\boldsymbol{u}(s))ds.$$

这样, 由 (i) 可知 $\boldsymbol{u}\in\mathcal{C}^1([0,T];\mathbb{R}^N)$, 且 $\boldsymbol{u}$ 是初值问题的唯一解. □

我们将在后面 (定理 3.11-1) 看到, 利用 Ascoli-Arzelà 定理 (定理 3.10-1), 可以在对映射 g 的较弱的假设下, 建立起在定理 3.8-1 中考察过的初值问题的解的存在性, 但只是在 "充分小的时段中".

注 (1) 当空间 $\mathcal{C}([0,T];\mathbb{R}^N)$ 赋以通常的 sup 范数, 即 $\boldsymbol{v}\to\sup_{0\leqslant t\leqslant T}\|\boldsymbol{v}(t)\|$, 其中当 $\|\cdot\|$ 是 $\mathbb{R}^N$ 上任一范数时, 除非 $\gamma<\frac{1}{T}$ (即对给定的 $T>0$, γ 充分小), 否则上述证明中引入的映射 $\boldsymbol{F}$ 未必是压缩映射. 但是, 关于这个范数, 当因子个数足够多时, 复合映射 $\boldsymbol{F}\circ\boldsymbol{F}\circ\cdots\circ\boldsymbol{F}$ 是一个压缩映射, 从而由习题 3.7-2, 可以导出解的存在性.

(2) Cauchy-Lipschitz 定理的 "局部" 形式将在习题 3.8-1 中给出.

(3) 当 $N=1,\boldsymbol{u}(0)=0$ 时, 函数 $\boldsymbol{u}\in\mathcal{C}([0,T])$ 满足的积分方程 (见上述证明中的 (i)) 是第一类非线性 Volterra 积分方程的一种特例, 这种积分方程一般形式为

$$\boldsymbol{u}(t)=\int_0^t h(t,s,u(s))ds, \quad 0\leqslant t\leqslant T,$$

其中 $h \in \mathcal{C}([0,T] \times \mathbb{R} \times \mathbb{R})$ 是一给定的函数.

(4) 定理 3.8-1 中存在性的结果与 $\mathbb{R}^N$ 中范数的选择无关 (改变 $\mathbb{R}^N$ 中的范数只影响到常数 γ).

下面关于线性常微分方程组的解的存在性和唯一性的结果是定理 3.8-1 的一个直接推论. 记号 $\mathbb{M}^N$ 表示所有实 $N \times N$ 阵组成的空间.

定理 3.8-2 设对某 $T > 0$, 给定一个矩阵场 $\boldsymbol{A} \in \mathcal{C}([0,T];\mathbb{M}^N)$ 和向量场 $\boldsymbol{b} \in \mathcal{C}([0,T];\mathbb{R}^N)$. 给定向量 $\boldsymbol{u}_0 \in \mathbb{R}^N$. 则初值问题

$$\begin{aligned}
&\boldsymbol{u}'(t) = \boldsymbol{A}(t)\boldsymbol{u}(t) + \boldsymbol{b}(t), \quad 0 \leqslant t \leqslant T, \\
&\boldsymbol{u}(0) = \boldsymbol{u}_0
\end{aligned}$$

有且仅有一个解 $\boldsymbol{u} \in \mathcal{C}^1([0,T];\mathbb{R}^N)$.

注 当矩阵场 $\boldsymbol{A}$ 取常值, 且当 $\boldsymbol{u}(0) = \boldsymbol{0}$ 时, 可用矩阵指数函数 (习题 3.6-1) 给出显式解.

Cauchy-Lipschitz 定理的一个值得注意的应用是单摆在垂直面上的运动. 一个单摆, 或确切地说, 一个 “理想摆” 是一根长为 ℓ 的无重量的刚性棒, 其一端是可以绕着点 O 自由转动, 质量 m 集中于另一端. 在单摆于垂直面上运动的附加假设下, 在时刻 t 其位置可完全由通过 O 点垂直向下的轴与单摆自身的夹角 $\theta(t)$ 确定 (图 3.8-1).

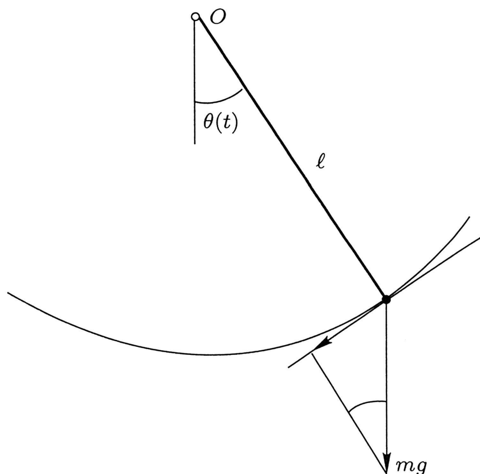


图 3.8-1 单摆

运动的方程可以利用 Newton 定律, 在任何时刻 t, 考察在以 O 为中心, ℓ 为半径的圆圈的切向量上的投影得到. 直接导出 $-mg\sin\theta(t) = m\ell\theta''(t)$, 这里常数 $g > 0$ 表

示重力加速度. 这样, 单摆运动服从非线性二阶微分方程, 即对任何时刻 t,

$$\theta''(t) = -\frac{g}{\ell}\sin\theta(t),$$

这个方程称之为*单摆方程*.

我们现在来证明, 加上初始条件 (即简单地取定初始角 θ_0 和初始速度 ω_0), 单摆方程有唯一解.

定理 3.8-3 *给定任意常数 θ_0 和 ω_0, 初值问题*

$$\theta''(t) = -\frac{g}{\ell}\sin\theta(t), \quad t \geqslant 0,$$
$$\theta(0) = \theta_0, \quad \theta'(0) = \omega_0$$

有且仅有一个解 $\theta \in \mathcal{C}^\infty([0,\infty[)$.

证明 定义向量场 $\boldsymbol{u} : [0,\infty[\to \mathbb{R}^2$ 为 $\boldsymbol{u}(t) = (u_i(t))_{i=1}^2$, 其中对所有的 $t \geqslant 0, u_1(t) := \theta(t), u_2(t) := \theta'(t)$. 这个向量场满足

$$\boldsymbol{u}'(t) = \boldsymbol{g}(t, \boldsymbol{u}(t)), \quad t \geqslant 0 \text{ 且 } \boldsymbol{u}'(0) = \boldsymbol{u}_0,$$

其中向量值函数 $\boldsymbol{g} : [0,\infty[\times\mathbb{R}^2 \to \mathbb{R}^2$ 定义为

$$\boldsymbol{g}(t,\boldsymbol{v}) := \left(v_2, -\frac{g}{\ell}\sin v_1\right), \quad (t,\boldsymbol{v}) \in [0,\infty[\times\mathbb{R}^2,$$

$\boldsymbol{u}_0 := (\theta_0, \omega_0)$. 上面两个一阶常微分方程组成的方程组即为 Cauchy-Lipschitz 定理 (定理 3.8-1) 中考虑过的形式. 此外, 对一切 $\boldsymbol{w}, \boldsymbol{v} \in \mathbb{R}^2$,

$$\begin{aligned}\|\boldsymbol{g}(t,\boldsymbol{w}) - \boldsymbol{g}(t,\boldsymbol{v})\|_1 &= |w_2 - v_2| + \frac{g}{\ell}|\sin w_1 - \sin v_1| \\ &\leqslant \max\left\{1, \frac{g}{\ell}\right\}\|\boldsymbol{w} - \boldsymbol{v}\|_1.\end{aligned}$$

因此, 可以对任意大的时间 T 用前面的定理, 从而得到存在唯一的解 $\boldsymbol{u} \in \mathcal{C}^1([0,\infty[;\mathbb{R}^2)$ (对任意 $t > 0, \boldsymbol{u}(t)$ 定义为对任何 $T \geqslant t$ 的解在 t 处的值; 由定理 3.8-1, $\boldsymbol{u}(t)$ 实际上是唯一确定的), 由此又可得到单摆方程的解 $\theta \in \mathcal{C}^2([0,\infty[)$ 的存在性与唯一性. 至于 $\theta \in \mathcal{C}^\infty([0,\infty[)$ 则可由单摆方程求导得到. □

对定理 3.8-3 的有趣的补充将在习题 3.8-2 中给出.

习题

3.8-1 利用类似于定理 3.8-1 的证明, 建立下述*局部形式的* Cauchy-Lipschitz 定理. 设 $\|\cdot\|$ 表示 $\mathbb{R}^N$ 上的任一范数. 给定 $T > 0, r > 0$ 和 $\boldsymbol{u}_0 \in \mathbb{R}^N$. 设有映射 $\boldsymbol{g} \in \mathcal{C}([0,T] \times \overline{B(\boldsymbol{u}_0;r)};\mathbb{R}^N)$, 其中 $B(\boldsymbol{u}_0;r) := \{\boldsymbol{v} \in \mathbb{R}^N; \|\boldsymbol{v} - \boldsymbol{u}_0\| < r\}$, 且存在常数 $\gamma > 0$, 使得对一切 $t \in [0,T]$ 和所有的 $\boldsymbol{w}, \boldsymbol{v} \in \overline{B(\boldsymbol{u}_0;r)}$,

$$\|\boldsymbol{g}(t,\boldsymbol{w}) - \boldsymbol{g}(t,\boldsymbol{v})\| \leqslant \gamma\|\boldsymbol{w} - \boldsymbol{v}\|.$$

则存在 $0<\tau\leqslant T$ 使得初值问题

$$\boldsymbol{u}'(t)=\boldsymbol{g}(t,\boldsymbol{u}(t)),\quad 0\leqslant t\leqslant\tau \text{ 且 } \boldsymbol{u}(0)=\boldsymbol{u}_0$$

有且仅有一个解 $\boldsymbol{u}\in\mathcal{C}^1([0,\tau];\mathbb{R}^N)$.

注意, 由中值定理 (定理 7.2-1), 任何 $\mathcal{C}^1$ 类映射 $\boldsymbol{g}$ 在点 $(0,\boldsymbol{u}_0)\in\mathbb{R}\times\mathbb{R}^N$ 的一个邻域中, 对某些 $T>0$ 和 $r>0$ 满足上述假设.

3.8-2 设 $\theta\in\mathcal{C}^\infty([0,\infty[)$ 是单摆方程对应于初始条件 $\theta(0)=0$ 和 $\theta'(0)=\omega_0>0$ 的解 (定理 3.8-3).

(1) 证明任何 $t>0$ 使得 $0\leqslant\tau\leqslant t$ 均有当 $\theta'(\tau)>0$ 时, 满足

$$t=\frac{1}{\omega_0}\int_0^{\theta(t)}\frac{d\psi}{\sqrt{1-\frac{4g}{\omega_0^2\ell}\sin^2\frac{\psi}{2}}}.$$

(2) 由 (1) 导出单摆会有下列三种可能类型的运动: 如果 $\omega_0>2\sqrt{\frac{g}{\ell}}$, 单摆会以周期变化的速度转动无限次 (即对所有的 $t>0,\theta'(t)>0$, 且 $\lim_{t\to\infty}\theta(t)=\infty$); 如果 $\omega_0=2\sqrt{\frac{g}{\ell}}$, 则对所有的 $t>0$, 有 $0<\theta(t)<\pi$ 且 $\theta'(t)>0,\lim_{t\to\infty}\theta(t)=\pi$; 如果 $\omega_0<2\sqrt{\frac{g}{\ell}}$, 则单摆在两个角 $-\alpha$ 和 α 间周期性地摆动, 这里角 $0<\alpha<\pi$ 和周期 $T>0$ 由下列两式给出[15].

$$\frac{4g}{\omega_0^2\ell}\sin^2\frac{\alpha}{2}=1,$$

$$\frac{T}{4}=\sqrt{\frac{\ell}{g}}\int_0^{\frac{\pi}{2}}\frac{d\varphi}{\sqrt{1-\sin^2\frac{\alpha}{2}\sin^2\varphi}}.$$

(3) 证明: 在 (2) 中考察的第三种情况下, 周期 T 也可展开为级数

$$T=2\pi\sqrt{\frac{\ell}{g}}\left(1+\frac{\alpha^2}{16}+\cdots\right).$$

3.9 Banach 不动点定理的应用: 非线性两点边值问题解的存在性

作为赋以 sup 范数的空间 $\mathcal{C}[a,b]$ 的完备性和 Banach 不动点定理的一个应用, 我们现在来对有界开区间 $I=]a,b[\subset\mathbb{R}$ 上一类特殊的非线性边值问题, 建立其经典解的存在性与唯一性 (对更一般的边值问题的推广, 见习题 3.9-1). "经典" 解是指在 I

[15] 在后一积分中令 $k=\sin\frac{\alpha}{2},\sin\varphi=t$, 则这个积分形为 $\int_0^{t_0}\frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$, 它是第一类椭圆积分的一个例子. 这类积分和形为 $\int_0^{t_0}\frac{\sqrt{1-k^2t^2}}{\sqrt{1-t^2}}dt$ 的第二类椭圆积分曾有很多研究的对象 (这两类积分均不能用初等函数计算), 涉及的著名数学家有 Leonhard Euler (1707—1783), Adrien-Marie Legendre (1752—1833), Carl Gustav Jacob Jacobi (1804—1851). 形容词 "椭圆" 表示椭圆上弧的长度恰好是用这类积分表示 (些时对应于第二类).

椭圆积分的详细讨论可见:

D. F. Lawden [1989]: Elliptic Functions and Applications. Applied Mathematical Sciences Series, Vol. 98, Springer, Heidelberg.

上二阶连续可微在 $\overline{I}$ 上连续的解, 它不同于 "弱" 解, 后者属于 $L^2(I)$, 且在分布意义下的导数也属于 $L^2(I)$ 的解 (将在第 6 章中介绍并讨论弱解). 不失一般性, 设 $I =]0,1[$.

定理 3.9-1 设 $I=]0,1[$, 函数 $f \in \mathcal{C}(\overline{I} \times \mathbb{R})$, 且存在常数 γ, 使得

$$0 \leqslant \gamma < 8 \quad 且 \quad |f(x,u) - f(x,v)| \leqslant \gamma |u - v|$$

对一切 $0 \leqslant x \leqslant 1, u, v \in \mathbb{R}$ 均成立, 又设 $\alpha, \beta \in \mathbb{R}$ 为两个常数. 则两点边值问题

$$\begin{aligned} &-u''(x) = f(x, u(x)), \quad 0 < x < 1, \\ &u(0) = \alpha, \quad u(1) = \beta \end{aligned}$$

有且仅有一个解 $u \in \mathcal{C}(\overline{I}) \cap \mathcal{C}^2(I)$.

证明 为清晰起见, 证明分为三部分. 注意, 在 (i) 和 (ii) 两部分中只需要假设 $f \in \mathcal{C}(\overline{I} \times \mathbb{R})$

(i) 如果 $u \in \mathcal{C}(\overline{I}) \cap \mathcal{C}^2(I)$ 是边值问题的解, 则 $u \in \mathcal{C}^2(\overline{I})$.

因为 $f \in \mathcal{C}(\overline{I} \times \mathbb{R}), u \in \mathcal{C}(\overline{I})$, 由关系式

$$u'(x) = u'\left(\frac{1}{2}\right) - \int_{\frac{1}{2}}^{x} f(t, u(t))dt, \quad 0 < x < 1,$$

可知函数 $u' \in \mathcal{C}(I)$ 可以延拓为 $\overline{I}$ 上的一个连续函数. 又由 Rolle 定理, 对每个 $0 < x < 1$, 存在 $\xi \in]0, x[$, 使得

$$\frac{u(x) - u(0)}{x} = u'(\xi),$$

因此, u 在 0 处可微, 且 $u'(0) = \lim_{\xi \to 0} u'(\xi)$; 类似可得 u 在 1 处可微, 且 $u'(1) = \lim_{\xi \to 1} u'(\xi)$. 因此, $u \in \mathcal{C}^1(\overline{I})$. 类似地, 由关系式 $-u''(x) = f(x, u(x)), 0 < x < 1$ 可得 $u \in \mathcal{C}^2(\overline{I})$.

(ii) 如果 $u \in \mathcal{C}^2(\overline{I})$ 是边值问题的解, 则 u 也是积分方程

$$u(x) = \alpha(1-x) + \beta x + \int_0^1 G(x,\xi) f(\xi, u(\xi)) d\xi, \quad 0 \leqslant x \leqslant 1$$

的解, 这里 $G \in \mathcal{C}(\overline{I} \times \overline{I})$ 定义为

$$G(x,\xi) := \begin{cases} \xi(1-x), & 0 \leqslant \xi \leqslant x \leqslant 1, \\ x(1-\xi), & 0 \leqslant x < \xi \leqslant 1. \end{cases}$$

反之, 如果 $u \in \mathcal{C}(\overline{I})$ 是上述积分方程的解, 则 $u \in \mathcal{C}^2(\overline{I})$, 且 u 是边值问题的解.

设 $u \in \mathcal{C}^2(\bar{I})$ 是边值问题的解. 则由函数 G 和 u 的定义,

$$\int_0^1 G(x,\xi)f(\xi,u(\xi))d\xi = -(1-x)\int_0^x \xi u''(\xi)d\xi - x\int_x^1(1-\xi)u''(\xi)d\xi$$
$$= u(x) - \alpha(1-x) - \beta x, \quad 0 \leqslant x \leqslant 1.$$

反之, 设 $u \in \mathcal{C}(\bar{I})$ 是积分方程的解. 首先, 显然有 $u(0)=\alpha, u(1)=\beta$. 其次, 连续两次求导可知, u 在 $[0,1]$ 中二阶连续可微, 且

$$u'(x) = -\alpha + \beta - \int_0^x \xi f(\xi,u(\xi))d\xi + \int_x^1 (1-\xi)f(\xi,u(\xi))d\xi, \quad 0 \leqslant x \leqslant 1,$$
$$-u''(x) = xf(x,u(x)) + (1-x)f(x,u(x)) = f(x,u(x)), \quad 0 \leqslant x \leqslant 1.$$

(iii) 在空间 $\mathcal{C}(\bar{I})$ 上赋以 sup 范数 $\|\cdot\|$, 得到一个 Banach 空间 (定理 3.2-2). 定义映射 $F:\mathcal{C}(\bar{I}) \to \mathcal{C}(\bar{I})$ 为

$$F(v)(x) = \alpha(1-x) + \beta x + \int_0^1 G(x,\xi)f(\xi,u(\xi))d\xi, \quad 0 \leqslant x \leqslant 1.$$

则 F 是一个压缩映射. 这是因为当 $0 \leqslant x,\xi \leqslant 1$ 时, $G(x,\xi)\geqslant 0$, 且对所有的 $0\leqslant x\leqslant 1$,

$$\int_0^1 G(x,\xi)d\xi \leqslant \frac{1}{8},$$

对任意的 $v,w \in \mathcal{C}(\bar{I})$, 有不等式

$$|(F(w)-F(v))(x)| \leqslant \int_0^1 G(x,\xi)|f(\xi,w(\xi)) - f(\xi,v(\xi))|d\xi$$
$$\leqslant \left(\sup_{0\leqslant x\leqslant 1}\int_0^1 G(x,\xi)d\xi\right)\sup_{0\leqslant\xi\leqslant 1}|f(\xi,w(\xi)) - f(\xi,v(\xi))|$$
$$\leqslant \frac{\gamma}{8}\sup_{0\leqslant\xi\leqslant 1}|w(\xi)-v(\xi)|, \quad 0 \leqslant x \leqslant 1,$$

而假设 $\gamma < 8$, 所以 F 是压缩映射. 于是, F 有且仅有一个不动点 $u \in \mathcal{C}(\bar{I})$. 由 (ii), $u \in \mathcal{C}^2(\bar{I})$, 且是边值问题的唯一解. □

在 (ii) 中出现的函数 G 是与微分算子 $u \in \{v \in \mathcal{C}^2(\bar{I}); v(0)=v(1)=0\} \to -u'' \in \mathcal{C}(\bar{I})$ 相应的 Green 函数. 即意为对任何函数 $f \in \mathcal{C}(\bar{I})$, 边值问题

$$-u''(x) = f(x), \quad 0 \leqslant x \leqslant 1, \quad u(0)=u(1)=0$$

的唯一解 $u \in \mathcal{C}^2(\bar{I})$ 可以表为 $u(x) = \int_0^1 G(x,\xi)f(\xi)d\xi$, $0 \leqslant x \leqslant 1$.

注 函数 $u \in \mathcal{C}(\bar{I})$ 满足的积分方程 (见上述证明的 (ii)) 是第一类非线性 Fredholm 积分方程的特例, 这类方程一般形式为

$$u(x) = \int_0^1 h(x,\xi,u(\xi))d\xi, \quad 0 \leqslant x \leqslant 1,$$

其中 $h \in \mathcal{C}(\bar{I}\times\mathbb{R}\times\mathbb{R})$ 为一给定的函数.

上述证明的关键是利用边值问题和积分方程的等价性 (见 (ii)), 把要求 $u \in \mathcal{C}^2(\bar{I})$ 用稍弱的要求 $u \in \mathcal{C}(\bar{I})$ 取代. 这一取代就转化为在空间 $\mathcal{C}(\bar{I})$ 上用 Banach 不动点定理. 可惜的是这种情况仅限于一维.

如果函数 f 关于第二个变量可微, 则第二个假设可取其等价形式

$$\left|\frac{\partial f}{\partial u}(x, v)\right| \leqslant \gamma < 8,$$

对一切 $(x, v) \in \bar{I} \times \mathbb{R}$ 成立.

实际上, 利用单调算子的理论 (习题 9.14-3), 在稍弱一点的假设下, 即存在常数 γ 使得当

$$\frac{\partial f}{\partial u}(x, v) \leqslant \gamma < \pi^2$$

对所有的 $(x, v) \in \bar{I} \times \mathbb{R}$ 成立时, 仍然可以建立定理 3.9-1 中两点边值问题解的存在性和唯一性, 这里出现 π^2 并不奇怪. 为说明这一点, 考察边值问题

$$-u''(x) = \pi^2 u(x), \quad 0 < x < 1 \text{ 且 } u(0) = 0,\ u(1) = \beta,$$

如果 $\beta = 0$, 则它有无穷多个解 $u : x \in [0, 1] \to u(x) = C \sin \pi x$, 其中 C 为任意常数, 当 $\beta \neq 0$ 时无解. 其原因是 π^2 为算子 $u \in \{v \in \mathcal{C}^2(\bar{I}); v(0) = v(1) = 0\} \to -u'' \in \mathcal{C}^0(\bar{I})$ 的特征值 (实际上是最小特征值), 当 $C \neq 0$ 时, 函数 u 是相应的特征函数.

注 当 $\frac{\partial f}{\partial u}(x, v) \leqslant 0$ 对所有的 $(x, v) \in \bar{I} \times \mathbb{R}$ 成立时, 利用 Ascoli-Arzelà 定理可以得到存在定理; 线性情况见习题 3.10-3[16].

习题

3.9-1 设 $I =]0, 1[$, 函数 $f \in \mathcal{C}(\bar{I} \times \mathbb{R} \times \mathbb{R})$, 且存在常数 $\gamma > 0$ 和 $\delta > 0$, 使得

$$|f(x, u, p) - f(x, v, q)| \leqslant \gamma |u - v| + \delta |p - q|$$

对一切 $x \in \bar{I}, u, v, p, q \in \mathbb{R}$ 成立. 利用与定理 3.9-1 所证明相似的方法, 证明: 如果 $\gamma + \delta$ 充分小, 则两点边值问题

$$-u''(x) = f(x, u(x), u'(x)), \quad 0 < x < 1 \quad \text{且} \quad u(0) = \alpha,\ u(1) = \beta$$

有且仅有一个解 $u \in \mathcal{C}^0(\bar{I}) \cap \mathcal{C}^2(I)$.

---

[16] 在非线性的情况下, 由于解的先验界, 问题可先归结为右边有界的情况; 参见:

P. G. Ciarlet; M. H. Schultz; R. S. Varga [1969]: Numerical methoels of high-order accuracy for nonliear boundary value problems V: Monotone operator theory. Numerische Mathematik **13**, 51–79.

3.10 Ascoli-Arzelà 定理

设 K 为紧距离空间, $\mathcal{C}(K)$ 是由所有的连续函数 $f: K \to \mathbb{R}$ 组成的空间, 在空间 $\mathcal{C}(K)$ 上赋以 sup 范数 $\|\cdot\|$, 即对所有的 $f \in \mathcal{C}(K)$,

$$\|f\| = \sup_{x \in K} |f(x)|.$$

这样 $\mathcal{C}(K)$ 是一个 Banach 空间 (定理 3.2-2).

下面的结果给出了 $(\mathcal{C}(K), \|\cdot\|)$ 中紧子集的基本特征.

定理 3.10-1 (Ascoli-Arzelà 定理[17]) 设 (K, d) 为紧距离空间. 则子集 $\mathcal{F} \subset \mathcal{C}(K)$ 是 $(\mathcal{C}(K), \|\cdot\|)$ 中的相对紧集当且仅当下列两个条件同时满足:

(a) 存在 M 使得

$$\|f\| \leqslant M$$

对所有的 $f \in \mathcal{F}$ 成立.

(b) 任意给定 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$, 使得对任意的 $x, y \in K$, 当 $d(x, y) < \delta(\varepsilon)$ 时,

$$|f(x) - f(y)| < \varepsilon$$

对所有的 $f \in \mathcal{F}$ 成立.

证明 如前, 在距离空间中, 用 $B(a; r)$ 表示以 a 为中心, $r > 0$ 为半径的 (开) 球, 用 diam A 表示子集 A 的直径.

(i) 设 $\mathcal{F}$ 为 $\mathcal{C}(K)$ 的子集, 使 $\overline{\mathcal{F}}$ 在 $\mathcal{C}(K)$ 中紧. 于是, $\overline{\mathcal{F}}$ 是 $\mathcal{C}(K)$ 中的有界集 (定理 1.13-1), 从而满足性质 (a).

任意给定 $\varepsilon > 0$, 并集 $\cup_{f \in \overline{\mathcal{F}}} B(f; \frac{\varepsilon}{3})$ 构成了紧集 $\overline{\mathcal{F}}$ 的一个开覆盖. 因此 (1.8 节) 存在有限个函数 $f_j \in \overline{\mathcal{F}} \subset \mathcal{C}(K), 1 \leqslant j \leqslant n$, 使得

$$\mathcal{F} \subset \overline{\mathcal{F}} \subset \bigcup_{j=1}^{n} B\left(f_j; \frac{\varepsilon}{3}\right).$$

因为 f_j 是一致连续的 (定理 1.13-2), 其个数有限, 因而存在 $\delta(\varepsilon) > 0$, 使得对任何 $x, y \in K$, 当 $d(x, y) < \delta(\varepsilon)$ 时,

$$|f_j(x) - f_j(y)| < \frac{\varepsilon}{3}$$

[17] C. Arzelà [1883]: Un'osservazione intorno alle serie di funzioni. Rendiconti delle Sessioni dell' Accademia Reale delle Scienze dell' Istituto di Bologna, 142–159.

C. Ascoli [1883]: Le curve limiti di una varietà data di curve. Atti della Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, *Matematiche e Naturali* **18**, 521–586.

对所有的 $1 \leqslant j \leqslant n$ 成立. 任意给定 $f \in \mathcal{F}$, 取 j_0 使得 $f \in B(f_{j_0}; \frac{\varepsilon}{3})$. 于是, 对任意的 $x, y \in K$, 当 $d(x,y) < \delta(\varepsilon)$ 时

$$|f(x) - f(y)| \leqslant |f(x) - f_{j_0}(x)| + |f_{j_0}(x) - f_{j_0}(y)| + |f_{j_0}(y) - f(y)| < \varepsilon.$$

因此性质 (b) 也满足.

(ii) 反之, 设 $\mathcal{F}$ 是 $\mathcal{C}(K)$ 的子集, 满足性质 (a), (b). 为证明 $\overline{\mathcal{F}}$ 在 $\mathcal{C}(K)$ 中紧, 只要证明 $\overline{\mathcal{F}}$ 是完备的, 且是准紧的 (定理 1.13-3). 因为 $\overline{\mathcal{F}}$ 在完备距离空间, 特别地, 在 Banach 空间 $\mathcal{C}(K)$ 中是一个闭子集, 从而是完备的 (定理 1.12-2(b)), 余下只要证明 $\overline{\mathcal{F}}$ 是准紧的, 这又等价于 $\mathcal{F}$ 自身是准紧的.

设给定 $\varepsilon > 0$. 首先, 由性质 (b) 可知, 存在 $\delta(\varepsilon) > 0$, 使得对任意的 $x, y \in K$, 当 $d(x,y) < \delta(\varepsilon)$ 时

$$|f(x) - f(y)| \leqslant \frac{\varepsilon}{3}$$

对一切 $f \in \mathcal{F}$ 成立. 另一方面, 因为 $K \subset \cup_{x \in K} B(x; \delta(\varepsilon))$ 且 K 是紧的, 所以存在有限个点 $x_\ell \in K$, $1 \leqslant \ell \leqslant p$, 使得

$$K \subset \bigcup_{\ell=1}^{p} B(x_\ell; \delta(\varepsilon)).$$

其次, 存在有限多个数 $y_m \in \mathbb{R}, 1 \leqslant m \leqslant q$, 使得

$$-M = y_1 < y_2 < \cdots < y_q = M$$

且

$$y_{m+1} - y_m < \frac{\varepsilon}{2}, \quad 1 \leqslant m \leqslant q-1,$$

其中 M 是性质 (b) 中出现的常数.

用 $\{\sigma_j; 1 \leqslant j \leqslant n\}$ 表示由所有从集合 $\{1, 2, \cdots, p\}$ 到集合 $\{1, 2, \cdots, q\}$ 的映射组成的有限集, $\mathcal{F}$ 的子集 A_j (可能为空集), $1 \leqslant j \leqslant n$, 定义为

$$A_j := \{f \in \mathcal{F}; |f(x_\ell) - y_{\sigma_j(\ell)}| < \frac{\varepsilon}{4}, \ 1 \leqslant \ell \leqslant p\}, \quad 1 \leqslant j \leqslant n.$$

我们断言

$$\mathcal{F} \subset \bigcup_{j=1}^{n} A_j \text{ 且 } \operatorname{diam} A_j \leqslant \varepsilon, \quad 1 \leqslant j \leqslant n,$$

由此证得 $\mathcal{F}$ 是准紧的.

为证明这个结论, 设 f 是 $\mathcal{F}$ 中的任意函数. 因为由性质 (a), 对 $1 \leqslant \ell \leqslant p$ 均有 $|f(x_\ell)| \leqslant M$, 故对每个 $\ell \in \{1, 2, \cdots, p\}$, 存在整数 $k = k(f, \ell) \in \{1, 2, \cdots, q\}$, 使得 $|f(x_\ell) - y_{k(f,\ell)}| \leqslant \frac{\varepsilon}{4}$ (由构造, $y_{m+1} - y_m < \frac{\varepsilon}{2}$, $1 \leqslant m \leqslant q-1$). 因而, 由映射 $\sigma_j (1 \leqslant j \leqslant n)$ 的定义, 存在整数 $j(f) \in \{1, \cdots, n\}$, 使得映射 $\sigma_{j(f)}$ 满足 $\sigma_{j(f)}(\ell) = k(f, \ell), 1 \leqslant \ell \leqslant p$, 或等价于使 $f \in A_{\sigma_{j(f)}}$. 因此, $\mathcal{F} \subset \cup_{j=1}^{n} A_j$.

任意给定整数 $j \in \{1,2,\cdots,n\}$, 设 $f,g \in A_j, x \in K$. 因为 $K \subset \cup_{\ell=1}^{p} B(x_\ell;\delta(\varepsilon))$, 故存在 $\ell = \ell(x) \in \{1,2,\cdots,p\}$, 使得 $x \in B(x_\ell;\delta(\varepsilon))$. 由 $\delta(\varepsilon)$ 的定义和集合 A_j 的定义可得

$$|f(x)-g(x)| \leqslant |f(x)-f(x_\ell)| + |f(x_\ell)-y_{\sigma_{j(f)}\ell}| \\ +|y_{\sigma_{j(f)}\ell}-g(x_\ell)| + |g(x_\ell)-g(x)| < \varepsilon.$$

因此, $\operatorname{diam} A_j \leqslant \varepsilon$. 证毕. □

满足定理 3.10-1 性质 (b) 的子集 $\mathcal{F} \subset \mathcal{C}(K)$ 称作等度连续的. 前缀“等度”反映了 $\delta(\varepsilon)$ 的选取不仅与 $x,y \in K$ 无关 (因为 K 是紧的, 每个函数 $f \in \mathcal{F}$ 均是一致连续的), 而且与 $f \in \mathcal{F}$ 无关.

由于这些定义, Ascoli-Arzelà 定理可以简述为: $\mathcal{C}(K)$ 中子集 $\mathcal{F}$ 的闭包为紧集的充要条件是 $\mathcal{F}$ 为有界且等度连续的.

在应用中 (如在习题 3.10-3 以及下一节处理的情况) 经常采用 Ascoli-Arzelà 定理的下列的推论 (其证明由定理 1.13-3 和 3.10-1 直接可得).

定理 3.10-2 (Ascoli-Arzelà 定理的推论) 设 K 为紧距离空间, $(f_n)_{n=0}^{\infty}$ 为函数 $f_n \in \mathcal{C}(K)$ 的序列, 满足下列性质:

(a) 存在常数 M, 对一切 $n \geqslant 0$, 有

$$\|f_n\| \leqslant M.$$

(b) 对任意给定的 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$, 使得对任意的 $x,y \in K$, 当 $d(x,y) < \delta(\varepsilon)$ 时,

$$|f_n(x)-f_n(y)| < \varepsilon$$

对一切 $n \geqslant 0$ 成立.

则存在子列 $(f_{\sigma(n)})_{n=0}^{\infty}$ 和函数 $f \in \mathcal{C}(K)$, 使得

$$\lim_{n\to\infty} \|f_{\sigma(n)} - f\| = 0.$$ □

容易明白如果空间 $\mathcal{C}(K)$ 被空间 $\mathcal{C}(K;\mathbb{R}^N)$ 取代, 定理 3.10-1 和 3.10-2 依然成立, 唯一的修正是 $|\cdot|$ 用 $\mathbb{R}^N$ 的某范数取代, $\|\cdot\|$ 用相应的 sup 范数取代 (只要讨论分量, 并逐次取 N 个子列). 实际上, 还可以在空间 $\mathcal{C}(K;Y)$ 上建立 Ascoli-Arzelà 定理, 这里 Y 是 Banach 空间, 见习题 3.10-1.

特别地, Ascoli-Arzelà 定理为证明两点边值问题的存在定理 (习题 3.10-3) 和常微分方程的存在定理 (3.11 节), 提供了一个有力的工具.

习题

3.10-1 证明 Ascoli-Arzelà 定理 (定理 3.10-1) 的下述推广成立. 设 (K,d) 为紧距离空间, $(Y,\|\cdot\|)$ 为 Banach 空间, 空间 $\mathcal{C}(K;Y)$ 赋以 sup 范数 (3.2 节). 证明子集 $\mathcal{F} \subset \mathcal{C}(K;Y)$

的闭包 $\overline{\mathcal{F}}$ 是紧集当且仅当满足以下两个性质:

(a) 对每个 $x \in X$, 集合 $\{f(x); x \in X\}$ 的闭包是 Y 的紧子集.

(b) 任意给定 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$, 使得对任意的 $x, y \in K$, 当 $d(x,y) < \delta(\varepsilon)$ 时, $\|f(x) - f(y)\| < \varepsilon$ 对一切 $f \in \mathcal{F}$ 成立.

3.10-2 设 K 为紧距离空间, 函数 $f_n \in \mathcal{C}(K)$ 的序列 $(f_n)_{n=1}^{\infty}$ 是等度连续的, 且点态收敛于函数 $f : K \to \mathbb{R}$. 证明 $f \in \mathcal{C}(K)$, 且 $(f_n)_{n=1}^{\infty}$ 一致收敛于 f.

3.10-3 设给定两个函数 $c \in \mathcal{C}[0,1]$ 和 $f \in \mathcal{C}[0,1]$, 本题的目标是在假设当 $0 \leqslant x \leqslant 1$ 时 $c(x) \geqslant 0$ 的条件下, 对于两点边值问题

$$-u''(x) + c(x)u(x) = f(x),\ 0 < x < 1 \text{ 且 } u(0) = u(1) = 0,$$

建立解 $u \in \mathcal{C}^2[0,1]$ 的存在性 (不失一般性, 这里假设 $u(0) = u(1) = 0$, 如果 $u(0) = \alpha$, $u(1) = \beta$ 且 $|\alpha| + |\beta| > 0$, 则只需引入新的未知函数 $x \in [0,1] \to u(x) - \alpha(1-x) - \beta x$ 即可). 这个方法是将 Ascoli-Arzelà 定理用于对这个边值问题的有限差分逼近时构造出来的一列函数上 (下面记作 $\widehat{u}_n$, 见问题 (7)). 注意, 利用 Banach 不动点定理, 定理 3.9-1 也导出了这个问题解的存在性, 只是假设不同于此. 那里要求 $|c(x)| \leqslant \gamma$, 其中常数 $\gamma < 8$.

任意给定整数 $n \geqslant 1$, 令 $h := \frac{1}{n+1}$. 逼近上述边值问题的有限差分法是寻找向量 $\boldsymbol{u}_h \in \mathbb{R}^n$, 要求满足线性系统 $\boldsymbol{A}_h \boldsymbol{u}_h = \boldsymbol{f}_h$, 其中 $n \times n$ 阵 $\boldsymbol{A}_h$ 和向量 $\boldsymbol{f}_h \in \mathbb{R}^n$ 定义为

$$\boldsymbol{A}_h := \frac{1}{h}\begin{pmatrix} 2+c_1h^2 & -1 & & & 0 \\ -1 & 2+c_2h^2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ 0 & -1 & 2+c_{n-1}h^2 & -1 \\ & & -1 & 2+c_nh^2 \end{pmatrix}, \quad \boldsymbol{f}_h := \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix},$$

这里 $c_i := c(ih), f_i := f(ih), 1 \leqslant i \leqslant n$. 这里的逼近就是用有限差分近似 $\frac{-u_{i-1}+2u_i-u_{i+1}}{h^2}, 1 \leqslant i \leqslant n$ 取代 $-u''(x_i)$, 其中 $u_0 = u_{n+1} = 0$.

(1) 证明对每个 $h = \frac{1}{n+1}$, 矩阵 $\boldsymbol{A}_h$ 具有以下性质: 如果向量 $\boldsymbol{v} = (v_i) \in \mathbb{R}^n$ 满足 $(\boldsymbol{A}_h\boldsymbol{v})_i \geqslant 0, 1 \leqslant i \leqslant n$, 则 $v_i \geqslant 0, 1 \leqslant i \leqslant n$.

(2) 由 (1) 证明, 对每个 $h = \frac{1}{n+1}$, 矩阵 $\boldsymbol{A}_h$ 是单调的, 即 $\boldsymbol{A}_h$ 可逆, 且 $(\boldsymbol{A}_h^{-1})_{ij} \geqslant 0$, $1 \leqslant i, j \leqslant n$.

(3) 记 $c(x) = 0, 0 \leqslant x \leqslant 1$ 时相应的矩阵 $\boldsymbol{A}_h$ 为 $\boldsymbol{A}_{oh}$, 证明对所有的 $h = \frac{1}{n+1}$, $\|\boldsymbol{A}_{oh}^{-1}\|_\infty \leqslant \frac{1}{8}$ (这里 $\|\boldsymbol{B}\|_\infty = \max_{1\leqslant i\leqslant n} \sum_{j=1}^{n} |(\boldsymbol{B})_{ij}|$, 见习题 2.9-1).

(4) 利用 (2) 证明对所有的 $h = \frac{1}{n+1}, \|\boldsymbol{A}_h^{-1}\|_\infty \leqslant \|\boldsymbol{A}_{oh}^{-1}\|_\infty$.

(5) 证明函数 $u \in \mathcal{C}^2[0,1]$ 是两点边值问题的解, 当且仅当 $u \in \mathcal{C}[0,1]$, 且 u 是积分方程

$$u(x) = \int_0^1 G(x,\xi)(-c(\xi)u(\xi) + f(\xi))d\xi, \quad 0 \leqslant x \leqslant 1,$$

的解, 其中函数 $G \in \mathcal{C}([0,1] \times [0,1])$ 定义为当 $0 \leqslant \xi \leqslant x \leqslant 1$ 时, $G(x,\xi) := \xi(1-x)$, 当 $0 \leqslant x < \xi \leqslant 1$ 时, $G(x,\xi) := x(1-\xi)$.

(6) 证明向量 $\boldsymbol{u}_h$ 是方程 $\boldsymbol{A}_h\boldsymbol{u}_h=\boldsymbol{b}_h$ 的解, 当且仅当其分量 $u_i, 1\leqslant i\leqslant n$ 满足方程 (它是问题 (5) 中的积分方程的离散形式)

$$u_i=h\sum_{j=1}^{n}G(ih,j)(-c_ju_j+f_j),\quad 1\leqslant i\leqslant n.$$

(7) 对每个 $h=\frac{1}{n+1}$, 连续函数 $\hat{u}_n:[0,1]\to\mathbb{R}$ 定义为: $\hat{u}_n(0)=\hat{u}_n(1)=0, \hat{u}_n(ih)=(\boldsymbol{u}_h)_i, 1\leqslant i\leqslant n$. 在区间 $[ih,(i+1)h]$ 上, $\hat{u}_n$ 为仿射, 这里 $1\leqslant i\leqslant n$. 证明存在与 n 无关的常数 M, 使得对一切 $n\geqslant 1$,

$$\sup_{0\leqslant x\leqslant 1}|\hat{u}_n(x)|\leqslant M$$

且序列 $(\hat{u}_n)_{n=1}^{\infty}$ 是等度连续的, 即给定 $\varepsilon>0$, 存在 $\delta(\varepsilon)>0$, 使得对任何 $x,y\in[0,1]$, 当 $|x-y|\leqslant\delta(\varepsilon)$ 时

$$|\hat{u}_n(x)-\hat{u}_n(y)|\leqslant\varepsilon$$

对所有的 $n\geqslant 1$ 成立.

提示: 为建立最后一个性质, 利用下述不等式的离散形式: 对任何函数 $\varphi\in\mathcal{C}^2[0,1]$, 均有

$$|\varphi'(x)|\leqslant|\varphi(1)-\varphi(0)|+\frac{1}{2}\sup_{0\leqslant\xi\leqslant 1}|\varphi''(\xi)|,\quad 0\leqslant x\leqslant 1.$$

(8) 利用 Ascoli-Arzelà 定理证明存在序列 $(\hat{u}_n)_{n=1}^{\infty}$ 的一个子列一致收敛于函数 $u\in\mathcal{C}[0,1]$. 证明 $u\in\mathcal{C}^2[0,1]$, 且 u 是两点边值问题的解.

(9) 证明实际上整个序列 $(\hat{u}_n)_{n=1}^{\infty}$ 一致收敛于 u.

(10) 由 (9) 证明: 在这个问题中考察的有限差分法在下面的意义下是收敛的, 即当 $n+1=\frac{1}{h}\to\infty$ 时

$$\max_{1\leqslant i\leqslant n}|u(ih)-u_i|\to 0.$$

证明如果解 u 具有某些光滑性质, 则这里的收敛性可以更为强化. 特别地, 证明当 $u\in\mathcal{C}^3[0,1]$ 时

$$\max_{1\leqslant i\leqslant n}|u(ih)-u_i|\leqslant\frac{2}{3}h\sup_{0\leqslant\xi\leqslant 1}\left|\frac{d^3u}{dx^3}(\xi)\right|,$$

当 $u\in\mathcal{C}^4[0,1]$ 时

$$\max_{1\leqslant i\leqslant n}|u(ih)-u_i|\leqslant\frac{h^2}{92}\sup_{0\leqslant\xi\leqslant 1}\left|\frac{d^4u}{dx^4}(\xi)\right|.$$

(11) 证明这个有限差分法的收敛阶数, 即当 $u\in\mathcal{C}^4[0,1]$ 时为 $O(h^2)$, 一般来说即便 u 具有更高的光滑性也不能增加.

3.10-4 下面, 空间 $\mathcal{C}[0,1]$ 赋以 sup 范数, 空间 $L^2[0,1]$ 赋以范数 $\|\cdot\|_{L^2[0,1]}$, G 是 $\mathcal{C}([0,1]\times[0,1])$ 中的给定函数.

(1) 对任意给定的函数 $v\in\mathcal{C}[0,1]$, 令

$$Av(x):=\int_0^1 G(x,\xi)v(\xi)d\xi,\quad 0\leqslant x\leqslant 1.$$

证明由这个关系式定义的函数 $Av\in\mathcal{C}[0,1]$, 这样定义的线性算子 $A:\mathcal{C}[0,1]\to\mathcal{C}[0,1]$ 是紧算子 (2.10 节).

(2) 对任意给定的函数 $v \in L^2[0,1]$, $Av(x), 0 \leqslant x \leqslant 1$, 如 (1) 所定义. 证明函数 $Av : [0,1] \to \mathbb{R}$ 是连续的, 并证明这样定义的线性算子 (仍记作) $A : L^2[0,1] \to \mathcal{C}[0,1]$ 和 $A : L^2[0,1] \to L^2[0,1]$ 均是紧的.

提示: 应用 Ascoli-Arzelà 定理.

(3) 证明: 如果对所有的 $(x,\xi) \in [0,1] \times [0,1]$, 均有 $G(x,\xi) = G(\xi,x)$. 证明算子 A 满足对一切 $v, w \in \mathcal{C}[0,1]$ 均有

$$\int_0^1 (Av(x))w(x)dx = \int_0^1 v(x)(Aw(x))dx.$$

注　(1) 对两点边值问题 $-u''(x) = f(x), 0 \leqslant x \leqslant 1$, 且 $u(0) = u(1)$ 的分析提供了这类算子 A 的一个例子, 这是因为其唯一解恰具形式 $u(x) = \int_0^1 G(x,\xi)f(\xi)d\xi$, 其中当 $0 \leqslant \xi \leqslant x \leqslant 1$ 时, $G(x,\xi) := \xi(1-x)$, 当 $0 \leqslant x < \xi \leqslant 1$ 时, $G(x,\xi) = x(1-\xi)$ (即将验证).

(2) 函数 G 仅属于空间 $L^2(]0,1[\times]0,1[)$ 的情况是习题 4.9-5 的讨论对象.

3.11　Ascoli-Arzelà 定理的应用: 非线性常微分方程解的存在性; Cauchy-Peano 定理; Euler 方法

利用 Banach *不动点定理*, 我们对于形为

$$\boldsymbol{u}'(t) = \boldsymbol{g}(t, \boldsymbol{u}(t)),\ 0 \leqslant t \leqslant T \text{ 且 } \boldsymbol{u}(0) = \boldsymbol{u}_0$$

的*常微分方程组的初值问题*, 建立了 (定理 3.8-1) 解 $\boldsymbol{u} \in \mathcal{C}^1([0,1];\mathbb{R}^N)$ 的存在性和唯一性, 这里出现在右边的映射 $\boldsymbol{g} : (t, \boldsymbol{v}) \in [0,T] \times \mathbb{R}^N \to \boldsymbol{g}(t,\boldsymbol{v}) \in \mathbb{R}^N$, 在 $[0,T] \times \mathbb{R}^N$ 上连续, 对第一个变量 $t \in [0,T]$ 一致地满足关于第二个变量 $\boldsymbol{v}$ 的 Lipschitz 条件.

我们将用 Ascoli-Arzelà 定理证明这个方程组在较弱的假设下依然有解, 这里仅假设有某个 $T > 0$ 和 $r > 0$, 映射 g 在形为 $\{(t,\boldsymbol{v}) \in \mathbb{R} \times \mathbb{R}^N; 0 \leqslant t \leqslant T, \|\boldsymbol{v} - \boldsymbol{u}_0\| \leqslant r\}$ 上连续. 当然, 为增加一般性是需要付出代价的.

首先, 这个结果仅提供*局部*存在性, 即便右边是光滑的, 且定义于所有的 $(t,\boldsymbol{v}) \in \mathbb{R} \times \mathbb{R}^N$, 解也仅在 $t \in [0,\tau]$ 上存在, 其中 $\tau > 0$ 可能任意地小. 例如, 考察初值问题:

$$u'(t) = (u(t))^2, \quad t \geqslant 0 \text{ 且 } u(0) = u_0.$$

则由 $u(t) = \frac{u_0}{1-u_0 t}$ 给出 (唯一) 的解当 $u_0 \leqslant 0$ 时对一切 $t \geqslant 0$ 有定义, 而当 $u_0 > 0$ 时仅对 $t \in [0,\tau]$ 有意义, 其中 $\tau > 0$ 是满足 $\tau < \frac{1}{u_0}$ 的任意数. 于是, 当 $u_0 \to +\infty$ 时解仅定义于区间 $[0,\tau]$, 而此区间将会变得任意地小 (因为当 t 从左侧趋于 $\frac{1}{u_0}$ 时此解会破裂).

其次, 可能会出现*非唯一性*. 例如, 考察初值问题

$$u'(t) = 3(u(t))^{\frac{3}{2}}, 0 \leqslant t \text{ 且 } u(0) = u_0.$$

这里右边出现的函数 $v \in \mathbb{R} \to v^{\frac{3}{2}} \in \mathbb{R}$ 是连续的, 但在 $v = 0$ 处并不满足 Lipschitz 条件. 这样, 由 Cauchy-Lipschitz 定理不能导出解的存在唯一性, 然而, 下面的定理 3.11-1 却给出了局部存在性. 更特别地, 如果 $u_0 \neq 0$, 则上述问题在 $t \geqslant 0$ 上有唯一解 $u(t) = (t + 2u_0^{\frac{1}{3}})^3$, 但若 $u_0 = 0$, 则有无穷多个解, 它们为

对所有的 $t \geqslant 0$, $u(t) = 0$.

对所有的 $t \geqslant 0$, $u(t) = t^3$.

对 $0 \leqslant t \leqslant t_0$, $u(t) = 0$, 对 $t \geqslant t_0$, $u(t) = (t - t_0)^3$,

其中 $t_0 > 0$ 可以任意选取 (注意, 当 $u_0 \neq 0$ 时解的局部存在性与唯一性还可以由 Cauchy-Lipschitz 定理的 "局部" 形式给出, 见习题 3.8-1).

定理 3.11-1 (Cauchy-Peano 定理) 设 $\|\cdot\|$ 为 $\mathbb{R}^N$ 上的任意一个范数, 给定 $T > 0, r > 0$ 和 $u_0 \in \mathbb{R}^N$, 又设给定映射 $\boldsymbol{g} \in \mathcal{C}([0, T] \times \overline{B(\boldsymbol{u}_0; r)}; \mathbb{R}^N)$, 其中 $B(\boldsymbol{u}_0; r) := \{\boldsymbol{v} \in \mathbb{R}^N; \|\boldsymbol{v} - \boldsymbol{u}_0\| \leqslant r\}$. 则存在 $\tau, 0 < \tau \leqslant T$, 使得初值问题

$$\boldsymbol{u}'(t) = \boldsymbol{g}(t, \boldsymbol{u}(t)),\ 0 \leqslant t \leqslant \tau \text{ 且 } \boldsymbol{u}(0) = \boldsymbol{u}_0$$

至少有一个解 $\boldsymbol{u} \in \mathcal{C}^1([0, \tau]; \mathbb{R}^N)$.

证明 (i) 令

$$\begin{aligned} M &:= \sup\{\|\boldsymbol{g}(t, \boldsymbol{v})\|;\ (t, \boldsymbol{v}) \in [0, T] \times \overline{B(\boldsymbol{u}_0; r)}\}, \\ \tau &:= \min\left\{\frac{r}{M}, T\right\}. \end{aligned}$$

在定理 3.8-1 的证明中已经知道, 只要对积分方程

$$\boldsymbol{u}(t) = \boldsymbol{u}_0 + \int_0^t \boldsymbol{g}(s, \boldsymbol{u}(s))ds, \quad 0 \leqslant t \leqslant \tau,$$

建立解 $\boldsymbol{u} \in \mathcal{C}^0([0, \tau]; \mathbb{R}^N)$ 的存在性与唯一性.

(ii) 任意给定整数 $n \geqslant 1$, 令 $h := \frac{\tau}{n}$, $t_i := ih$, $0 \leqslant i \leqslant m$, 于是 $0 = t_0 \leqslant t_i \leqslant \tau \leqslant T$, $0 \leqslant i \leqslant n$. 这样, 逼近这个初值问题的最简单的*差分方法*是用

$$\frac{\boldsymbol{u}_{i+1} - \boldsymbol{u}_i}{h} = \boldsymbol{g}(t_i, \boldsymbol{u}_i), \quad 0 \leqslant i \leqslant n - 1,$$

递推地定义向量 $\boldsymbol{u}_i \in \mathbb{R}^N$, $1 \leqslant i \leqslant n$. 当然, 我们首先必须验证 $\boldsymbol{u}_i \in \overline{B(\boldsymbol{u}_0; r)}, 1 \leqslant i \leqslant n - 1$ (否则, $(t_i, \boldsymbol{u}_i)$ 将不在映射 $\boldsymbol{g}$ 的定义域中, 无法定义 $\boldsymbol{u}_{i+1}$). 为此, 注意到

$$\|\boldsymbol{u}_1 - \boldsymbol{u}_0\| = h\|\boldsymbol{g}(t_0, \boldsymbol{u}_0)\| \leqslant hM \leqslant \tau M \leqslant r.$$

因此, 假设对整数 $i \in \{2, \cdots, m - 1\}$ 有 $\|\boldsymbol{u}_{i-1} - \boldsymbol{u}_0\| \leqslant (i - 1)hM \leqslant \tau M \leqslant r$, 于是

$$\begin{aligned} \|\boldsymbol{u}_i - \boldsymbol{u}_0\| &\leqslant \|\boldsymbol{u}_i - \boldsymbol{u}_{i-1}\| + \|\boldsymbol{u}_{i-1} - \boldsymbol{u}_0\| \\ &\leqslant hM + (i - 1)hM = ihM \leqslant \tau M \leqslant r. \end{aligned}$$

因此, 逐次地迭代 $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 是完全确定的.

(iii) 对每个整数 $n \geqslant 1$, 定义向量值函数 $\hat{\boldsymbol{u}}_n : [0, \tau] \to \mathbb{R}^N$ 如下:

$$\hat{\boldsymbol{u}}_n(t) := \boldsymbol{u}_i + \frac{(t - t_i)}{h}(\boldsymbol{u}_{i+1} - \boldsymbol{u}_i), \quad t_i \leqslant t \leqslant t_{i+1}, \quad 0 \leqslant i \leqslant n-1,$$

即 $\hat{\boldsymbol{u}}_n(t_i) = \boldsymbol{u}_i$, $0 \leqslant i \leqslant n$, 在 $[t_i, t_{i+1}]$ 上 $\hat{\boldsymbol{u}}_n$ 为仿射, $0 \leqslant i \leqslant n-1$. 因此, 当 $n \geqslant 1$ 时 $\hat{\boldsymbol{u}}_n \in \mathcal{C}^0([0, \tau]; \mathbb{R}^N)$.

在空间 $\mathcal{C}^0([0, \tau]; \mathbb{R}^N)$ 上赋以通常的 sup 范数 $|||\cdot|||$, 序列 $(\hat{\boldsymbol{u}}_n)_{n=1}^{\infty}$ 有界, 这是因为

$$|||\hat{\boldsymbol{u}}_n||| = \sup_{0 \leqslant t \leqslant \tau} \|\hat{\boldsymbol{u}}_n(t)\| = \max_{1 \leqslant i \leqslant n} \|\boldsymbol{u}_i\|,$$
$$\|\boldsymbol{u}_i\| \leqslant \|\boldsymbol{u}_0\| + \|\boldsymbol{u}_i - \boldsymbol{u}_0\| \leqslant \|\boldsymbol{u}_0\| + r.$$

序列 $(\hat{\boldsymbol{u}}_n)_{n=1}^{\infty}$ 也是等度连续的, 这是因为对每个 $i \in \{0, 1, \cdots, n-1\}$, 当 $t_i \leqslant t \leqslant t_{i+1}$ 时

$$\begin{aligned}\|\hat{\boldsymbol{u}}_n(t) - \hat{\boldsymbol{u}}_n(t_i)\| &= \|\hat{\boldsymbol{u}}_n(t) - \boldsymbol{u}_i\| \\ &\leqslant (t - t_i) \left\| \frac{\boldsymbol{u}_{i+1} - \boldsymbol{u}_i}{h} \right\| \leqslant (t - t_i)M,\end{aligned}$$

从而对所有的 $t, \tilde{t} \in [0, \tau]$,

$$\|\hat{\boldsymbol{u}}_n(t) - \hat{\boldsymbol{u}}_n(\tilde{t})\| \leqslant |t - \tilde{t}|M.$$

由 Ascoli-Arzelà 定理, 存在序列 $(\hat{\boldsymbol{u}}_n)_{n=1}^{\infty}$ 的子列 $(\hat{\boldsymbol{u}}_{\sigma(n)})_{n=1}^{\infty}$ 和映射 $\boldsymbol{u} \in \mathcal{C}^0([0, \tau]; \mathbb{R}^N)$, 使得当 $n \to \infty$ 时

$$\sup_{0 \leqslant t \leqslant \tau} \|\hat{\boldsymbol{u}}_{\sigma(n)}(t) - \boldsymbol{u}(t)\| \to 0.$$

(iv) 余下要证明 $\boldsymbol{u}$ 是 (i) 中积分方程的解. 为此, 注意到

$$\begin{aligned}\boldsymbol{u}_{i+1} &= \boldsymbol{u}_0 + h(\boldsymbol{g}(t_0, u_0) + \boldsymbol{g}(t_1, u_1) + \cdots + \boldsymbol{g}(t_i, u_i)) \\ &= \boldsymbol{u}_0 + \int_0^{t_i} \boldsymbol{g}_n(s) ds, \quad 0 \leqslant i \leqslant n-1,\end{aligned}$$

其中分段常值映射 $\boldsymbol{g}_n : [0, \tau] \to \mathbb{R}^N$ 定义为

$$\boldsymbol{g}_n(s) := \boldsymbol{g}(t_i, \boldsymbol{u}_i), \quad t_i \leqslant s \leqslant t_{i+1}, \quad 0 \leqslant i \leqslant n.$$

再注意到在每个区间 $[t_i, t_{i+1}]$ 上常值映射的积分是一个仿射映射, 由此, 对每个整数 $n \geqslant 1$, 映射 $\hat{\boldsymbol{u}}_n \in \mathcal{C}^0([0, T]; \mathbb{R}^N)$ 由

$$\hat{\boldsymbol{u}}(t) = \boldsymbol{u} + \int_{0.}^{t} \boldsymbol{g}_n(s) ds, \quad 0 \leqslant t \leqslant \tau$$

给出.

由 (iv) 中的极限 $\boldsymbol{u}$ 和映射 $s \in [0,\tau] \to \boldsymbol{g}(s,\boldsymbol{u}(s))$ 的一致连续性, 利用 "(ε,δ) 语言", 可由当 $n \to \infty$ 时 $|||\hat{\boldsymbol{u}}_{\sigma(n)} - \boldsymbol{u}||| \to 0$ 直接证得当 $n \to \infty$ 时

$$\sup_{0 \leqslant s \leqslant \tau} \|\boldsymbol{g}_{\sigma(n)}(s) - \boldsymbol{g}(s,\boldsymbol{u}(s))\| \to 0,$$

由此可得当 $n \to \infty$ 时

$$\sup_{0 \leqslant t \leqslant \tau} \left\| \int_0^t \boldsymbol{g}_{\sigma(n)}(s)ds - \int_0^t \boldsymbol{g}(s,\boldsymbol{u}(s))ds \right\| \to 0.$$

综上即得

$$\boldsymbol{u}(t) = \boldsymbol{u}_0 + \int_0^t \boldsymbol{g}(s,\boldsymbol{u}(s))ds, \quad 0 \leqslant t \leqslant \tau,$$

证毕. □

在 (ii) 中给出的有限差分方法即构成了逼近常微分方程初值问题的 Euler *方法*.

如果从某种意义上可以建立起初值问题解的唯一性, 上面的证明说明在空间 $(\mathcal{C}^0([0,\tau];\mathbb{R}^N), |||\cdot|||)$ 中整个序列 $(\boldsymbol{u}_n)_{n=1}^{\infty}$ 收敛于 u, 这样我们又得到了 Euler *方法的收敛性*.

注 对于 Cauchy-Lipschitz 定理 (定理 3.8-1), 可以不费力地证明, 只要逐字逐句将 $\mathbb{R}^N$ 用任意的 Banach 空间 X 取代, 定理结论依然成立 (连续映射 $[0,T] \to X$ 的积分曾定义于 3.3 节), 但是在这种更一般的情况下, Cauchy-Peano 定理未必成立[18].

[18] J. Dieudonné [1950]: Deux exemples singuliers d'équations différentielles. Acta Scientiarum Mathematicarum B (Szeged) 12, 38–40.

第 4 章　内积空间和 Hilbert 空间

引言

在无限维赋范向量空间中, 内积空间特别是 Hilbert 空间即完备的内积空间, 如作为其原型的空间 ℓ^2 和 $L^2(\Omega)$ (4.2 节), 性质是最好的.

它们之所以引人关注的基本原因是其范数是由内积 (熟知的 $\mathbb{R}^n$ 中数量积的推广) 定义的, 因而具有 Euclid 范数的许多性质. 由此, n 维 Euclid 空间的许多 "几何" 性质在这些空间成立. 诸如 Cauchy-Schwarz-Bunyakovskiĭ 不等式和平行四边形法则 (4.1 节), 基本的投影定理 (定理 4.3-1), 向量的直交性 (4.5 节), 或当空间完备时任何元素可以用空间的赋范直交基以 Fourier 级数表示 (定理 4.9-1) 的可能. 这种可能性在下面以基本的例子作说明, 如经典的 (或三角的) Fourier 级数, Legendre、Laguerre 或 Hermite 多项式 (4.8 节).

利用投影定理, 我们还给出了关于线性系统最小二乘解的存在性的简洁证明 (定理 4.4-1).

Hilbert 空间引人瞩目的另一个基本原因是任何这样的空间可利用特殊的线性等距同构于其对偶空间. 这就是 Hilbert 空间中基本的 F. Riesz 表示定理的内容 (定理 4.6-1). 这个定理具有许多深刻的应用, 例如, 可以给出 Hilbert 空间上 Hahn-Banach 定理的简单证明, 可以给出连续线性算子的伴随算子的直接定义 (4.7 节). 值得注意的是, 引入类似的赋范向量空间上对偶算子的概念却需要用到选择公理 (利用 Hahn-Banach 延拓定理, 参见第 5 章).

本章还包括紧自伴算子谱理论的详尽的处理 (4.10 和 4.11 节). 特别地, 谱定理 (定理 4.11-1) 将成为第 6 章中二阶椭圆型边值问题中特征值问题的分析基础. 注意, 这是本书仅有的涉及谱论的地方, 因为对任意的赋范向量空间上相应问题的处理

已超出本书的范围.

4.1 内积空间和 Hilbert 空间; Cauchy-Schwarz-Bunyakovskiĭ 不等式; 平行四边形法则

设 X 为实向量空间. X 上的*内积*是指函数 $(\cdot,\cdot): X \times X \to \mathbb{R}$, 它满足对任何 $x, y, z \in X$ 和任何 $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned}(\alpha x + \beta y, z) &= \alpha(x, z) + \beta(y, z),\\ (x, \alpha y + \beta z) &= \alpha(x, y) + \beta(x, z),\\ (x, y) &= (y, x),\\ (x, x) &\geqslant 0 \quad \text{且当 } (x, x) = 0 \text{ 时 } x = 0.\end{aligned}$$

换言之, 实向量空间上的内积是一个双线性型 (即对两个变量中的每一个都是线性的函数; 参见 2.11 节), 它是对称的 (第三个性质; 注意, 由第一和第三个性质显然可得第二个性质), 而且是正定的 (第四个性质).

*实内积空间*是指偶对 $(X, (\cdot,\cdot))$, 其中 X 为实向量空间, $(\cdot,\cdot)$ 为 X 上的内积.

再设 X 为复向量空间. X 上的*内积*是指复值函数 $(\cdot,\cdot): X \times X \to \mathbb{C}$, 它满足对任何 $x, y, z \in X$ 和任何 $\alpha, \beta \in \mathbb{C}$ (记号 $\bar{\alpha}$ 表示 $\alpha \in \mathbb{C}$ 的复共轭):

$$\begin{aligned}(\alpha x + \beta y, z) &= \alpha(x, z) + \beta(y, z),\\ (x, \alpha y + \beta z) &= \bar{\alpha}(x, y) + \bar{\beta}(x, z),\\ (x, y) &= \overline{(y, x)},\\ (x, x) &\geqslant 0 \quad \text{且当 } (x, x) = 0 \text{ 时 } x = 0.\end{aligned}$$

换言之, 复向量空间上的内积是一个 Hermite 型 (第一、第二和第三个性质; 注意, 第二个性质也显然可由第一和第三个性质导出), 而且它是正定的 (第四个性质). 因此, 复向量空间上的内积关于第一个变量是线性的 (第一个性质), 关于第二个变量是半线性的 (第二个性质). 由此, 复向量空间上的内积有时称作*半双线性的* (sesquilinear, 前缀 “sesqui” 表示 “一个半”).

*复内积空间*是指偶对 $(X, (\cdot,\cdot))$, 其中 X 为复向量空间, $(\cdot,\cdot)$ 为 X 上的内积.

设 $\mathbb{K}$ 表示域 $\mathbb{R}$ 或域 $\mathbb{C}$. 所谓*内积空间*是指实内积空间 ($\mathbb{K} = \mathbb{R}$) 或复内积空间 ($\mathbb{K} = \mathbb{C}$).

在下面的定理中建立的基本不等式 (参见 (a)) 在内积空间理论中常被用到. 作为第一个推论, 由此可导出内积空间也是赋范向量空间 (参见 (b)), 内积是其两个变量的连续函数 (参见 (c)).

定理 4.1-1 (Cauchy-Schwarz-Bunyakovskiĭ 不等式[1]) 设 $(X,(\cdot,\cdot))$ 为实或复内积空间 ($\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$).

(a) Cauchy-Schwarz-Bunyakovskiĭ 不等式成立. 即对任何 $x,y\in X$:

$$|(x,y)|\leqslant\sqrt{(x,x)}\sqrt{(y,y)}.$$

(b) 函数

$$\|\cdot\|:x\in X\to\|x\|:=\sqrt{(x,x)}\in\mathbb{R}$$

是 X 上的一个范数, 而且, 对任何 $x\in X$,

$$\|x\|=\sup_{y\neq 0}\frac{|(x,y)|}{\|y\|}.$$

(c) 映射

$$(\cdot,\cdot):X\times X\to\mathbb{K}$$

关于 $X\times X$ 上的乘积拓扑是连续的, 其中 X 上的拓扑由 (b) 定义的范数导出.

证明 先设 $\mathbb{K}=\mathbb{R}$. 给定两个向量 $x,y\in X$, 其中 $y\neq 0$ (因此 $(y,y)>0$), 实二次多项式

$$p:t\in\mathbb{R}\to p(t):=(x+ty,x+ty)=(x,x)+2t(x,y)+t^2(y,y)$$

满足对任何 $t\in\mathbb{R},p(t)\geqslant 0$. 由此, 特别地有

$$p\left(-\frac{(x,y)}{(y,y)}\right)=(x,x)-\frac{|(x,y)|^2}{(y,y)}\geqslant 0,$$

这就证得当 $\mathbb{K}=\mathbb{R}$ 时 Cauchy-Schwarz-Bunyakovskiĭ 不等式成立 (如果 $y=0$, 相应的关系式为 $0=0$, 不等式成立).

再设 $\mathbb{K}=\mathbb{C}$ 给出两个向量 $x,y\in X$, 其中 $y\neq 0$, 考察复值函数

$$p:z\in\mathbb{C}\to p(z):=(x+zy,x+zy)=(x,x)+z(y,x)+\bar{z}(x,y)+z\bar{z}(y,y),$$

---

[1] 这个不等式首先在有限维情况下获得证明, 见:

A. L. Cauchy [1821]: Cours d'Analyse de l'Ecole Royale Polytechnique. de Bure, Paris.

见文献:

R. E. Bradley; C. E. Sandifer [2009]: Cauchy's Cours d'Analyse-An Annotated Translation. Springer, Heidelberg.

其后, 被推广为积分形式, 见:

V. Bunyakovskiĭ [1859]: Sur quelques inégalités concernant les intégrales aux différences finies, Mémoires de l'Académie des Sciences de Saint-Petersbourg. Tème Série, Tome 1, No. 9, 1–18.

推广到如这里所述的一般的内积空间, 见:

H. A. Schwarz [1885] Über ein Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung. Acta Societatis Scientiarum Fennicae 15, 315–362.

对任何 $z \in \mathbb{C}$, 满足 $p(z) \geqslant 0$. 从而, 特别地有

$$p\left(-\frac{(x,y)}{(y,y)}\right) = (x,x) - \frac{(x,y)\overline{(x,y)}}{(y,y)} \geqslant 0,$$

由此, Cauchy-Schwarz-Bunyakovskiĭ 不等式对 $\mathbb{K} = \mathbb{C}$ 也成立. 这就证得 (a).

为证 (b) 中定义的函数 $\|\cdot\| : X \to \mathbb{R}$ 满足三角不等式, 可以利用当 $\mathbb{K} = \mathbb{R}$ 时的等式

$$\|x+y\|^2 = \|x\|^2 + 2(x,y) + \|y\|^2$$

和 $\mathbb{K} = \mathbb{C}$ 的等式

$$\|x+y\|^2 = \|x\|^2 + 2\mathrm{Re}(x,y) + \|y\|^2,$$

并结合 Cauchy-Schwarz-Bunyakovskiĭ 不等式, 即得在两个情况下均有

$$\|x+y\|^2 \leqslant \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

因此, 函数 $\|\cdot\|$ 是一个范数 (范数另两个性质显然成立). 又由 Cauchy-Schwarz-Bunyakovskiĭ 不等式可得对任何非零 $x \in X$,

$$\|x\| = \frac{(x,x)}{\|x\|} \leqslant \sup_{y\neq 0} \frac{|(x,y)|}{\|y\|} \leqslant \|x\|,$$

所以 $\|x\| = \sup\limits_{y\neq 0} \dfrac{|(x,y)|}{\|y\|}$.

内积的连续性是因为对任何 $x, y, x_0, y_0 \in X$, 成立等式

$$(x,y) - (x_0,y_0) = (x - x_0, y_0) + (x_0, y - y_0) + (x - x_0, y - y_0),$$

再利用 Cauchy-Schwarz-Bunyakovskiĭ 不等式即得. □

注 (1) 如果函数 $(\cdot,\cdot) : X \times X \to \mathbb{R}$ 满足除了当 $(x,x) = 0$ 时必有 $x = 0$ 以外内积的所有性质 (即第四个性质仅为对任何 $x \in X, (x,x) \geqslant 0$), 则 Cauchy-Schwarz-Bunyakovskiĭ 不等式依然成立. 为说明这一点, 只要注意到前述证明包含了 $(y,y) > 0$ 的情况, 从而也包含了 $(x,x) > 0$ 的情况. 余下的当 $(x,x) = (y,y) = 0$ 时的情况, 则由当 $\mathbb{K} = \mathbb{R}$ 时

$$p(-(x,y)) = -2(x,y)^2 \geqslant 0$$

和当 $\mathbb{K} = \mathbb{C}$ 时

$$p(-(x,y)) = -2|(x,y)|^2 \geqslant 0$$

可得 $(x,y) = 0$. 故而此时 Cauchy-Schwarz-Bunyakovskiĭ 不等式依然成立 (它归化为 $0 = 0$).

(2) 由定理 4.1-1 的证明可见, Cauchy-Schwarz-Bunyakovskiĭ 不等式中等号成立的充要条件是 x 与 y 线性相关.

读者在下文中将会看到, 作为赋范向量空间, 对内积空间 $(X, (\cdot,\cdot))$ 总是赋以由定理 4.1-1(b) 所定义的范数, 称之为由内积 $(\cdot,\cdot)$ 导出的范数.

对此, 以下两个定理将给出两点说明, 它们给出了这个特殊的对于内积空间的范数所具有的两个基本性质.

定理 4.1-2 (平行四边形法则) 设 $(X, (\cdot,\cdot))$ 为实或复的内积空间, 则对任何 $x, y \in X$,

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

上述平行四边形法则可得内积空间是一致凸的 (2.17 节).

证明 平行四边形法则可由当 $\mathbb{K} = \mathbb{R}$ 时的等式

$$\|x \pm y\|^2 = \|x\|^2 \pm 2(x, y) + \|y\|^2$$

和当 $\mathbb{K} = \mathbb{C}$ 时的等式

$$\|x \pm y\|^2 = \|x\|^2 \pm 2\mathrm{Re}(x, y) + \|y\|^2$$

直接导出. 平行四边形公式又可表示为对任何 $x, y \in X$,

$$\left\|\frac{x+y}{2}\right\|^2 = \frac{1}{2}\|x\|^2 + \frac{1}{2}\|y\|^2 - \frac{1}{4}\|x-y\|^2.$$

因此, 当 $\|x\| = \|y\| = 1$ 且 $\|x-y\| \geqslant \varepsilon > 0$ 时

$$\left\|\frac{x+y}{2}\right\| \leqslant 1 - \delta(\varepsilon),$$

其中 $\delta(\varepsilon) := 1 - \sqrt{1 - \frac{\varepsilon^2}{4}} > 0$. 因此, 实或复的内积空间是一致凸的. □

当内积空间中两个向量 x 与 y 满足当 $(x, y) = 0$ 时, 称 x 与 y 为直交的. 当 x 与 y 直交时, 上述证明开始时出现的等式化为

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

由于它是直角三角形中一个熟知的性质在一般的内积空间的推广, 这个等式也称为 Pythagoras 定理[2)].

在定理 4.1-2 中建立的等式是 $\mathbb{R}^2$ 中平行四边形的一条熟知的性质 (图 4.1-1) 在任意内积空间的推广, 因而称之为 "平行四边形法则". 注意, 平行四边形法则有一个有趣的逆命题, 它说明这一法则实际上是内积空间的特征 (习题 4.1-3).

现在我们来导出内积空间上连续线性算子的算子范数的另一个特征, 它也可以表示为一个上确界. 仍用 $\mathcal{L}(X)$ 表示由赋范向量空间 X 到其自身的所有连续线性算子组成的空间 (2.9 节).

2) 这个命名是为纪念古希腊的哲学家 Pythagoras, 他约于公元前 520 年第一个对于三角形证明了这个等式 (实际上, 公元前 1500 年左右巴比伦人就知道了这个等式).

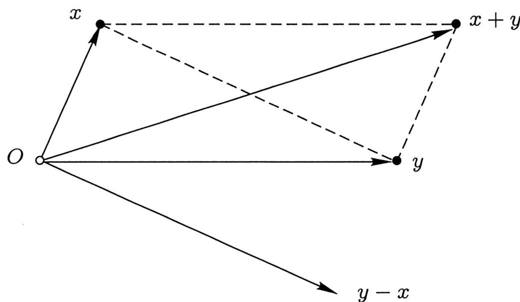


图 4.1-1 (平行四边形法则) $\mathbb{R}^2$ 中平行四边形两条对角线长度的平方和等于其四条边长的平方和

定理 4.1-3 设 $(X,(\cdot,\cdot))$ 为实或复内积空间, 则对任何算子 $A\in\mathcal{L}(X)$ 算子范数 $\|A\|=\sup_{x\neq 0}\frac{\|Ax\|}{\|x\|}$ 也可由

$$\|A\|=\sup_{\begin{cases}x\neq 0\\ y\neq 0\end{cases}}\frac{|(Ax,y)|}{\|x\|\|y\|}$$

给出.

证明 给定 $A\in\mathcal{L}(X)$. 由 Cauchy-Schwarz-Bunyakovskiĭ 不等式,

$$|(Ax,y)|\leqslant\|Ax\|\|y\|\leqslant\|A\|\|x\|\|y\|,$$

因此

$$\sup_{\begin{cases}x\neq 0\\ y\neq 0\end{cases}}\frac{|(Ax,y)|}{\|x\|\|y\|}\leqslant\|A\|.$$

设 $x\in X$ 满足 $Ax\neq 0$ (从而 $x\neq 0$). 于是

$$\begin{aligned}\frac{\|Ax\|^2}{\|x\|^2}=\frac{(Ax,Ax)}{\|x\|\|x\|}&=\frac{\|Ax\|}{\|x\|}\frac{(Ax,Ax)}{\|x\|\|Ax\|}\\&\leqslant\frac{\|Ax\|}{\|x\|}\sup_{\begin{cases}x\neq 0\\ y\neq 0\end{cases}}\frac{|(Ax,y)|}{\|x\|\|y\|},\end{aligned}$$

从而

$$\frac{\|Ax\|}{\|x\|}\leqslant\sup_{\begin{cases}x\neq 0\\ y\neq 0\end{cases}}\frac{|(Ax,y)|}{\|x\|\|y\|}.$$

显然, 最后一个不等式当 $Ax=0$ 但 $x\neq 0$ 时依然成立. 因此,

$$\|A\|=\sup_{x\neq 0}\frac{\|Ax\|}{\|x\|}\leqslant\sup_{\begin{cases}x\neq 0\\ y\neq 0\end{cases}}\frac{|(Ax,y)|}{\|x\|\|y\|}.$$ □

内积空间 $(X,(\cdot,\cdot))$ 称为 Hilbert 空间[3], 是指它作为赋范向量空间是一个 Banach 空间, 即 X 关于由 $\|x\|=\sqrt{(x,x)}, x\in X$ 定义的范数 $\|\cdot\|$, 它是完备的.

任何不完备的内积空间 X, 利用相应的赋范向量空间完备化的方法 (定理 3.1-2), 可以恒同于一个 Banach 空间 $\tilde{X}$ 的稠密子集. 可以期望 $\tilde{X}$ 也是一个 Hilbert 空间, 其内积是由 X 的内积模一个线性等距的扩张.

定理 4.1-4 (内积空间完备化) 设 $(X,(\cdot,\cdot))$ 为 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$ 上的内积空间. 则相应的赋范空间 $(X,\|\cdot\|)$ 的完备化空间 $(\tilde{X},\|\cdot\|_{\tilde{X}})$ (定理 3.1-2) 是 $\mathbb{K}$ 上的 Hilbert 空间, 其内积 $(\cdot,\cdot)_{\tilde{X}}$ 满足对任何 $x,y\in X$,

$$(\sigma x,\sigma y)_{\tilde{X}}=(x,y),$$

这里的 σ 是由 X 到 $\tilde{X}$ 的一个稠密子空间上的线性等距算子 (定理 3.1-2).

证明 对任何 $\tilde{x}=\sigma x\in\sigma(X)$ 和 $\tilde{y}=\sigma y\in\sigma(X)$, 令

$$(\tilde{x},\tilde{y})_{\sigma(X)}:=(x,y).$$

显然, 由于 σ 是线性等距算子, 这样定义的由 $\sigma(X)\times\sigma(X)$ 到 $\mathbb{K}$ 的映射是向量空间 $\sigma(X)$ 上的一个内积. 对任何 $\tilde{x}\in\tilde{X}$ 和 $\tilde{y}\in\tilde{X}$, 取 $\tilde{x}_n\in\sigma(X), n\geqslant 0$ 和 $\tilde{y}_n\in\sigma(X), n\geqslant 0$, 使得当 $n\to\infty$ 时

$$\|\tilde{x}_n-\tilde{x}\|_{\tilde{X}}\to 0,\quad \|\tilde{y}_n-\tilde{y}\|_{\tilde{X}}\to 0$$

(由构造, $\sigma(X)$ 是 $(\tilde{X},\|\cdot\|_{\tilde{X}})$ 的一个稠密子空间; 参见定理 3.1-2), 令

$$\begin{aligned}
(\tilde{x},\tilde{y})_{\tilde{X}} &:= \lim_{n\to\infty}(\tilde{x}_n,\tilde{y}_n)_{\sigma(X)}\\
(\text{如果 } \mathbb{K}=\mathbb{R}) &= \tfrac{1}{4}\lim_{n\to\infty}(\|\tilde{x}_n+\tilde{y}_n\|^2-\|\tilde{x}_n-\tilde{y}_n\|^2)\\
(\text{如果 } \mathbb{K}=\mathbb{C}) &= \tfrac{1}{4}\lim_{n\to\infty}(\|\tilde{x}_n+\tilde{y}_n\|^2-\|\tilde{x}_n-\tilde{y}_n\|^2+i\|\tilde{x}_n+i\tilde{y}_n\|^2-i\|\tilde{x}_n-i\tilde{y}_n\|^2)
\end{aligned}$$

(因此 $\lim_{n\to\infty}(\tilde{x}_n,\tilde{y}_n)_{\sigma(X)}$ 只依赖 $\tilde{x}$ 和 $\tilde{y}$), 容易验证这样定义的由 $\tilde{X}\times\tilde{X}$ 到 $\mathbb{K}$ 上的映射是 $\tilde{X}$ 上的内积, 它满足对任何 $x,y\in X$,

$$(\sigma x,\sigma y)_{\tilde{X}}=(x,y).$$

□

注 另一个证明可由平行四边形法则的逆 (习题 4.1-3) 导出.

本章的大部分结果可毫无区分地同时适用于实或复的内积空间, 尽管为清晰起见有时分两种情况处理 (如定理 4.1-1 中 Cauchy-Schwarz-Bunyakovskiĭ 不等式的证明). 应当注意一些特殊的情况. 有些结果仅适用于一种情况 (如见习题 4.1-1, 它提供了一个只在复空间中成立, 在实空间中不成立的性质的例子), 有些需要对两种情况给出不同的证明 (如见习题 4.1-3 提出的平行四边形法则的逆).

[3] 这样命名是为了纪念 David Hilbert (1862—1943), 他在 20 世纪初详尽地研究了 Hilbert 空间的一些特例 (特别可见于 Reid [1970] 的 Hilbert 传记中 Hermann Weyl 的章节和 Dieudonné [1981, 第 5 章第 2 节]). 而 "抽象" 的 Hilbert 空间的观念 (即并非如 ℓ^2 或 $L^2(\Omega)$ 等特例) 则属于 John von Neumann (1903—1957), 他在 1929 年首先引入术语 "Hilbert 空间".

习题

4.1-1 (1) 设 $(X,(\cdot,\cdot))$ 为复内积空间, 线性算子 $A: X\to X$ 满足对任何 $x\in X, (Ax,x)=0$. 证明 $A=0$.

(2) 证明如果 $(X,(\cdot,\cdot))$ 为实内积空间, 则 (1) 的相应结论并不成立.

4.1-2 设 X 为 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$ 上的向量空间, $(\cdot,\cdot): X\times X\to\mathbb{K}$ 为一个映射, 它满足内积的除了第四条 (正定性) 外的所有性质, 证明: 映射 $(\cdot,\cdot)$ 完全由它在乘积空间 $X\times X$ 的对角线即 $X\times X$ 的子集 $\{(x,y)\in X\times X: x=y\}$ 上的限制所确定.

4.1-3 设 $(X,\|\cdot\|)$ 为 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$ 上的赋范向量空间, 其范数满足平行四边形公式, 即对任何 $x,y\in X$:

$$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2.$$

(1) 证明 X 也是个内积空间, 且对任何 $x\in X$, 其内积满足 $\|x\|=\sqrt{(x,x)}$.

提示: 验证当 $\mathbb{K}=\mathbb{R}$ 时

$$(x,y):=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2),$$

当 $\mathbb{K}=\mathbb{C}$ 时

$$(x,y):=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2+i\|x+iy\|^2-i\|x-iy\|^2),$$

即为相应的内积.

(2) 利用 (1) 给出定理 4.1-4 的另一个证明.

4.1-4 (1) 设 X 和 Y 为两个实内积空间, 映射 $A: X\to Y$ 满足

$$A(0)=0,\quad \|Ax-A\tilde{x}\|_Y=\|x-\tilde{x}\|_X$$

对任何 $x,\widetilde{x}\in X$ 成立. 证明 A 是线性算子 (显然是连续的).

(2) 当 X,Y 为复内积空间时, 上述结果是否依然成立.

注 $X=Y=\mathbb{R}^n$ 的特殊情况即熟知的 Mazur-Ulam 定理, 见习题 8.7-1.

4.1-5 设 X 为 Hilbert 空间, Y 为 X 的闭子空间. 证明商空间 X/Y (它是 Banach 空间, 见定理 3.6-5) 也是 Hilbert 空间.

4.1-6 (1) 证明 $\mathbb{R}^n$ 中的 Cauchy-Schwarz-Bunyakovskiĭ 不等式, 即对任何 $x_i,y_i\in\mathbb{R}, 1\leqslant i\leqslant n$

$$\sum_{i=1}^n|x_iy_i|\leqslant\left(\sum_{i=1}^n|x_i|^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^n|y_i|^2\right)^{\frac{1}{2}}$$

等价于[4] 算术 – 几何平均不等式 (问题 2.17-10), 即对任何 $x_i>0, 1\leqslant i\leqslant n$,

$$\left(\prod_{i=1}^n x_i\right)^{\frac{1}{n}}\leqslant\frac{1}{n}\sum_{i=1}^n x_i.$$

[4] Minghua Lin [2012]: The AM-GA inequality and CBS inequality are equivalent. The Mathematical Intelligencer 34, 6.

(2) 证明算术 - 几何平均不等式等价于[5] Bernoulli 不等式, 即对任何 $x>0$ 和 $n\geqslant 1$,

$$1+n(x-1)\leqslant x^n.$$

4.2 内积空间和 Hilbert 空间的例子; 空间 ℓ^2 和 $L^2(\Omega)$

在 $\mathbb{R}^n$ 上, 对任何 $\boldsymbol{x}=(x_i)_{i=1}^n, \boldsymbol{y}=(y_i)_{i=1}^n$ 定义

$$\boldsymbol{x}\cdot\boldsymbol{y}:=\sum_{i=1}^n x_i y_i,$$

赋以 Euclid 内积也称之为数量积, 在 $\mathbb{C}^n$ 上, 对任何 $\boldsymbol{x}=(x_i)_{i=1}^n, \boldsymbol{y}=(y_i)_{i=1}^n$ 定义

$$\boldsymbol{x}\cdot\boldsymbol{y}:=\sum_{i=1}^n x_i \overline{y_i}$$

赋以 Hermite 内积, 它们提供了实和复 Hilbert 空间的最简单的例子. 于是, 由这个内积导出的范数为对 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 中任何向量 $\boldsymbol{x}=(x_i)_{i=1}^n$,

$$|\boldsymbol{x}|:=(\boldsymbol{x}\cdot\boldsymbol{x})^{\frac{1}{2}}=\left(\sum_{i=1}^n |x_i|^2\right)^{\frac{1}{2}}$$

赋以 Euclid 内积的空间 $\mathbb{R}^n$ 通常称为 n 维 Euclid 空间. 更一般地, 给定 n 阶正定的对称阵, 或相应地给定 n 阶 Hermite 阵 $\boldsymbol{A}=(a_{ij})$ 在 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 上赋以定义为

$$(\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{A}}:=\sum_{i,j=1}^n a_{ij}x_i y_i \quad \text{或} \quad (\boldsymbol{x},\boldsymbol{y})_{\boldsymbol{A}}:=\sum_{i,j=1}^n a_{ij}x_i \overline{y_i}$$

的内积, 则 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 也成为 Hilbert 空间.

在任何有限维向量空间上, 可以类似地定义内积.

另一类实或复的 Hilbert 空间是实或复的 $m\times n$ 阵组成的向量空间, 对任何 $m\times n$ 阵 $\boldsymbol{A}=(a_{ij}), \boldsymbol{B}=(b_{ij})$, 当 $\mathbb{K}=\mathbb{R}$ 时, 定义矩阵内积为

$$\boldsymbol{A}:\boldsymbol{B}:=\sum_{i=1}^m\sum_{j=1}^n a_{ij}b_{ij},$$

当 $\mathbb{K}=\mathbb{C}$ 时, 定义矩阵内积为

$$\boldsymbol{A}:\boldsymbol{B}:=\sum_{i=1}^m\sum_{j=1}^n a_{ij}\bar{b}_{ij}.$$

[5] L. Maligranda [2012]: The AM-GM inequality is equivalent to the Bernoulli inequality. The Mathematical Intelligencer 34, 1–2.

由这个内积导出的范数 $\|\cdot\|_F$ 定义为

$$\|\boldsymbol{A}\|_F := (\boldsymbol{A}:\boldsymbol{A})^{\frac{1}{2}} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{\frac{1}{2}},$$

其中 $m\times n$ 阵 $\boldsymbol{A}=(a_{ij})$, 称之为 Frobenius 范数.

上述 Hilbert 空间均是可分空间 (因为它们都是有限维的; 见定理 2.7-1(b)).

实或复的空间 ℓ^2 (即在 2.4 节中引入的空间 $\ell^p, 1\leqslant p\leqslant\infty$ 在当 $p=2$ 时的特例) 由满足 $\sum_{i=1}^{\infty}|x_i|^2<\infty$ 的数 $x_i\in\mathbb{K}$ 的无限序列 $x=(x_i)_{i=1}^{\infty}$ 的全体所组成. 当 $\mathbb{K}=\mathbb{R}$ 时, 对 $x=(x_i)_{i=1}^{\infty}, y=(y_i)_{i=1}^{\infty}\in\ell^2$, 赋以定义为

$$(x,y) := \sum_{i=1}^{\infty}x_iy_i$$

的内积, 当 $\mathbb{K}=\mathbb{C}$ 时, 对 $x=(x_i)_{i=1}^{\infty}, y=(y_i)_{i=1}^{\infty}\in\ell^2$, 赋以定义为

$$(x,y) := \sum_{i=1}^{\infty}x_i\bar{y}_i$$

的内积, 这样对任何 $x=(x_i)_{i=1}^{\infty}\in\ell^2$, 导出的范数定义为

$$\|x\| := \sqrt{(x,x)} = \left(\sum_{i=1}^{\infty}|x_i|^2\right)^{\frac{1}{2}},$$

关于这个范数, ℓ^2 是可分的 (定理 2.4-2(b)), 且是完备的 (定理 3.4-1), 因而空间 ℓ^2 提供了无限维实或复的可分 Hilbert 空间的例子. 注意, 相应的 Cauchy-Schwarz-Bunyakovskiĭ 不等式是序列的 Hölder 不等式 (定理 2.4-1(a)) 当 $p=q=2$ 时的特例, 相应的三角不等式是序列的 Minkovski 不等式 (定理 2.4-1(b)) 当 $p=2$ 时的特例.

实空间 $L^2(\Omega)$ (在 2.5 节中引入的空间 $L^p(\Omega), 1\leqslant p\leqslant\infty$ 当 $p=2$ 时的特例) 由所有满足 $\int_\Omega|f(x)|^2dx<\infty$ 的可测函数 $f:\Omega\to[-\infty,\infty]$ 的等价类组成, 其中 Ω 是 $\mathbb{R}^n$ 的任一开子集, 对任何 $f,g\in L^2(\Omega)$, 定义

$$(f,g) := \int_\Omega f(x)g(x)dx,$$

便在 $L^2(\Omega)$ 上赋以内积, 其上相应的范数为对任何 $f\in L^2(\Omega)$,

$$\|f\|_{L^2(\Omega)} := \left(\int_\Omega|f(x)|^2dx\right)^{\frac{1}{2}}.$$

关于这个范数, $L^2(\Omega)$ 是可分的 (定理 2.5-4(a)), 而且是完备的 (定理 3.4-2), 因而空间 $L^2(\Omega)$ 提供了无限维可分的实 Hilbert 空间的另一个例子. 注意, 相应的 Cauchy-Schwarz-Bunyakovskiĭ 不等式是函数的 Hölder 不等式 (定理 2.5-1(a)) 当 $p=q=2$

时的特例, 相应的三角不等式是函数的 Minkowski 不等式 (定理 2.5-1(b)) 当 $p=2$ 时的特例.

我们同样可以定义复空间

$$L^2(\Omega;\mathbb{C}) := \{f:\Omega\to\mathbb{C};\ \mathrm{Re}\, f \text{ 和 } \mathrm{Im}\, f \text{ 均可测, 且 } |f|^2 \in L^1(\Omega)\}.$$

容易证明这是一个无限维可分复 Hilbert 空间, 其上的内积定义为对任何 $f,g\in L^2(\Omega;\mathbb{C})$,

$$(f,g) := \int_\Omega f(x)\overline{g(x)}dx.$$

空间 $\mathcal{C}(\overline{\Omega})$ 提供了不完备的实内积空间 (从而必定是无限的) 的例子, 其中 Ω 是 $\mathbb{R}^N$ 的有界开子集, 其内积即空间 $L^2(\Omega)$ 的内积; 实际上, 空间 $\mathcal{C}(\overline{\Omega})$ 关于范数 $\|\cdot\|_{L^2(\Omega)}$ 的完备化空间即更大的空间 $L^2(\Omega)$ (定理 2.5-3 或定理 2.6-2).

空间 ℓ^2 和 $L^2(\Omega)$ 是无限维可分 Hilbert 空间较为基本的例子, 其他基本的例子如 Sobolev 空间 $H^m(\Omega)$ 和 $H_0^m(\Omega)$ 将在今后提供 (第 6 章).

今后还将看到, ℓ^2 实际上是这些空间中典型的例子, 因为借助于保持内积不变的线性双射, 任何无限维可分 Hilbert 空间可恒同于空间 ℓ^2 (定理 4.9-4).

习题

4.2-1 对两个非零向量 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{C}^n$, 可以定义其夹角为方程

$$\cos\theta(\boldsymbol{x},\boldsymbol{y}) = \frac{\mathrm{Re}(\boldsymbol{x}\cdot\boldsymbol{y})}{|\boldsymbol{x}||\boldsymbol{y}|}$$

的唯一解 $\theta(\boldsymbol{x},\boldsymbol{y})\in[0,\pi]$ (这个定义推广了两个非零向量 $\boldsymbol{x}$, $\boldsymbol{y}\in\mathbb{R}^n$ 的夹角概念), 或定义为方程

$$\cos\varphi(\boldsymbol{x},\boldsymbol{y}) = \frac{|\boldsymbol{x}\cdot\boldsymbol{y}|}{|\boldsymbol{x}||\boldsymbol{y}|}$$

的唯一解 $\varphi(\boldsymbol{x},\boldsymbol{y})\in\left[0,\frac{\pi}{2}\right]$.

注 当 $\boldsymbol{x}=i\boldsymbol{y}$ 时 $\theta(\boldsymbol{x},\boldsymbol{y})=\frac{\pi}{2}$, 而 $\varphi(\boldsymbol{x},\boldsymbol{y})=0$.

(1) 证明: 对任何非零向量 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathbb{C}^n, \theta(\boldsymbol{x},\boldsymbol{z})\leqslant\theta(\boldsymbol{x},\boldsymbol{y})+\theta(\boldsymbol{y},\boldsymbol{z})$.

(2) 利用 (1) 证明: 对任何非零向量 $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in\mathbb{C}^n, \varphi(\boldsymbol{x},\boldsymbol{z})\leqslant\varphi(\boldsymbol{x},\boldsymbol{y})+\varphi(\boldsymbol{y},\boldsymbol{z})$.

4.3 投影定理

贯穿于 Hilbert 空间理论中的下述结果是十分基本的. 特别地, 它是其他几个基本结果的关键所在. 例如, 直接和定理 (定理 4.5-2), Hilbert 空间上的 F. Riesz 表示定理 (定理 4.6-1) 或凸集上二次泛函的极小化定理 (定理 6.1-1). 其几何解释 (命名的由来) 将在证明后讨论.

定理 4.3-1 (投影定理) 设 Z 为实 ($\mathbb{K}=\mathbb{R}$) 或复 ($\mathbb{K}=\mathbb{C}$) 的内积空间 $(X,(\cdot,\cdot))$ 的非空完备凸子集.

(a) 对任意给定的 $x\in X$, 存在唯一的元 $Px\in Z$, 满足

$$\|x-Px\|=\inf_{z\in Z}\|x-z\|,$$

其中 $\|\cdot\|$ 是由内积 $(\cdot,\cdot)$ 导出的范数 (定理 4.1-1(c)).

(b) 在 (a) 中出现的唯一的元素 $Px\in Z$ 满足当 $\mathbb{K}=\mathbb{R}$ 时对所有的 $z\in Z$,

$$(Px-x,z-Px)\geqslant 0,$$

当 $\mathbb{K}=\mathbb{C}$ 时, 对所有的 $z\in Z$

$$\mathrm{Re}(Px-x,z-Px)\geqslant 0,$$

反之, 如果元素 $y\in Z$ 满足当 $\mathbb{K}=\mathbb{R}$ 时, 对所有的 $z\in Z$,

$$(y-x,z-y)\geqslant 0,$$

当 $\mathbb{K}=\mathbb{C}$ 时, 对所有的 $z\in Z$,

$$\mathrm{Re}(y-x,z-y)\geqslant 0,$$

则 $y=Px$.

(c) 在 (a) 中定义的映射 $P:X\to Z$ 满足对所有的 $x_1,x_2\in X$,

$$\|Px_1-Px_2\|\leqslant\|x_1-x_2\|.$$

因此 P 是 Lipschitz 连续映射, 其 Lipschitz 常数为 1.

(d) 设子集 Z 是 X 的一个完备子空间. 则 (a) 中的元素 Px 满足对所有的 $z\in Z$,

$$(Px-x,z)=0.$$

反之, 若元素 $y\in Z$ 满足对所有的 $z\in Z$,

$$(y-x,z)=0,$$

则 $y=Px$.

(e) 映射 $P:X\to Z$ 为线性的充要条件是子集 Z 为 X 的子空间, 此时, 若 $Z\neq\{0\}$, 则

$$\|P\|_{\mathcal{L}(X;Z)}=1.$$

证明 (i) 如果 $x \in Z$, 则 $Px = x$. 如果 $x \bar{\in} Z$. 由假设, 集 Z 非空, 从而确定了数 $\delta := \inf_{z\in Z}\|x-z\|$. 实际上 $\delta > 0$, 这是因为若 $\delta = 0$, 则 $x \in \overline{Z}$, 又由假设 Z 是闭的 (完备子集是闭的; 见定理 1.12-2(a)), 故而 $x \in Z$, 此为矛盾. 设 $y_n \in Z, n \geqslant 0$, 使得

$$\|x-y_n\| \underset{n\to\infty}{\longrightarrow} \delta = \inf_{z\in Z}\|x-z\| > 0.$$

由平行四边形法则 (定理 4.1-2) 可知, 对任何 $m, n \geqslant 0$,

$$\begin{aligned}\|y_m-y_n\|^2 &= \|(x-y_m)-(x-y_n)\|^2\\ &= 2\|x-y_m\|^2+2\|x-y_n\|^2-\|2x-(y_n+y_m)\|^2\\ &= 2\|x-y_m\|^2+2\|x-y_n\|^2-4\left\|x-\frac{y_n+y_m}{2}\right\|^2.\end{aligned}$$

由假设, 集合 Z 为凸的, 从而 $\frac{y_m+y_n}{2} \in Z$; 因此

$$\left\|x-\frac{y_m+y_n}{2}\right\| \geqslant \delta^2,$$

由此可知, 对任何 $m, n \geqslant 0$, 有

$$0 \leqslant \|y_m-y_n\|^2 \leqslant 2\|x-y_m\|^2+2\|x-y_n\|^2-4\delta^2.$$

因为 $\|x-y_m\| \underset{m\to\infty}{\longrightarrow} \delta, \|x-y_n\| \underset{n\to\infty}{\longrightarrow} \delta$, 故而序列 $(y_n)_{n=0}^{\infty}$ 为 Cauchy 序列. 由假设, 集合 Z 是完备的, 所以存在 $y \in Z$ 使得 $y_n \underset{n\to\infty}{\longrightarrow} y$. 再利用范数的连续性 (定理 2.2-5), 得到

$$\|x-y\| = \lim_{n\to\infty}\|x-y_n\| = \delta = \inf_{z\in Z}\|x-z\|.$$

为证明这样的元素 $y \in Z$ 是唯一的, 设 $y_0 \in Z$ 和 $y_1 \in Z$ 满足

$$\delta = \|x-y_0\| = \|x-y_1\|.$$

对任意 $k \geqslant 0$, 定义 $y_{2k} := y_0, y_{2k+1} := y_1$ 得到序列 $(y_n)_{n=0}^{\infty}$, 显然满足 $\|x-y_n\| \underset{n\to\infty}{\longrightarrow} \delta$. 由与上述同样的论证可得这个序列是收敛的. 因为赋范向量空间中收敛序列的极限是唯一的, 因此 $y_0 = y_1$. 这就证得 (a).

(ii) 先设 $\mathbb{K} = \mathbb{R}$. 如果 $x \in Z$, 此时 $Px - x = 0$, 欲证的不等式成立. 如果 $x \bar{\in} Z$, 令 $y := Px \in Z$, 给定 $z \in Z$. 因为对于 $0 \leqslant \theta \leqslant 1, y + \theta(z-y) \in Z$ (由假设; Z 是凸的). 由 $y = Px$ 的定义 (见 (a)) 可得对于 $0 \leqslant \theta \leqslant 1$,

$$\begin{aligned}\|x-y\|^2 &\leqslant \|x-(y+\theta(z-y))\|^2\\ &= \|x-y\|^2-2\theta(x-y,z-y)+\theta^2\|z-y\|^2.\end{aligned}$$

因此, 对任何 $0 \leqslant \theta \leqslant 1$,

$$0 \leqslant 2\theta(y-x,z-y)+\theta^2\|z-y\|^2,$$

从而 $(y-x,z-y)\geqslant 0$.

反之, 设元素 $y\in Z$ 满足对任何 $z\in Z,(y-x,z-y)\geqslant 0$. 于是, 对任何 $z\in Z$,

$$\begin{aligned}\|x-z\|^2&=\|x-y+y-z\|^2=\|x-y\|^2+2(y-x,z-y)+\|z-y\|^2\\&\geqslant\|x-y\|^2,\end{aligned}$$

由此, $y=Px$. 如果 $\mathbb{K}=\mathbb{C}$, 则由关系式

$$\|x-(y+\theta(z-y))\|^2=\|x-y\|^2-2\theta\mathrm{Re}(x-y,z-y)+\theta^2\|z-y\|^2,$$
$$\|x-z\|^2=\|x-y\|^2+2\mathrm{Re}(y-x,z-y)+\|z-y\|^2$$

可知相应的结论成立. 这就证得 (b).

(iii) 先设 $\mathbb{K}=\mathbb{R}$. 由 (b) 可得对任何 $x_1,x_2\in X$, 因为 $Px_2\in Z$, 所以

$$(Px_1-x_1,Px_2-Px_1)\geqslant 0,$$

因为 $Px_1\in Z$, 所以

$$(x_2-Px_2,Px_2-Px_1)\geqslant 0,$$

从而

$$(Px_1-Px_2,Px_2-Px_1)+(x_2-x_1,Px_2-Px_1)\geqslant 0.$$

由此, 结合 Cauchy-Schwarz-Bunyakovskiĭ 不等式即得

$$\|Px_1-Px_2\|^2\leqslant(x_2-x_1,Px_2-Px_1)\leqslant\|x_2-x_1\|\|Px_2-Px_1\|.$$

这就得到了要证明的不等式.

如果 $\mathbb{K}=\mathbb{C}$, 则由不等式

$$\mathrm{Re}(Px_1-x_1,Px_2-Px_1)\geqslant 0,\quad \mathrm{Re}(x_2-Px_2,Px_2-Px_1)\geqslant 0,$$

可得

$$\|Px_1-Px_2\|^2=\mathrm{Re}(Px_1-Px_2,Px_1-Px_2)\leqslant\mathrm{Re}(x_2-x_1,Px_2-Px_1),$$

所以同样的结论成立, 这就证得 (c).

(iv) 设 Z 为 X 的完备子空间, 且先设 $\mathbb{K}=\mathbb{R}$. 如果 $x\in Z$, 此时 $Px-x=0$, 欲证的不等式成立. 如果 $x\bar{\in} Z$, 给定 $z\in Z$. 因为对任何 $\theta\in\mathbb{R},Px+\theta z\in Z$ (这里假设 Z 为子空间), 由 (b) 中的不等式可得对任何 $\theta\in\mathbb{R}$,

$$(Px-x,Px+\theta z-Px)=\theta(Px-x,z)\geqslant 0.$$

因此, $(Px-x,z)=0$.

反之, 设 $y \in Z$ 满足对任何 $z \in Z$, 有 $(y-x,z)=0$, 于是, 特别地有 $(y-x,y)=0$. 因此, 对任何 $z \in Z$, 有

$$(y-x,z-y)=0 \geqslant 0.$$

由 (b) 即得 $y=Px$.

再设 $\mathbb{K}=\mathbb{C}$, 给定 $z \in Z$. 因为对任何 $\theta \in \mathbb{R}, Px+\theta z \in Z$, 由 (b) 所得的不等式, 对任何 $\theta \in \mathbb{R}$,

$$\operatorname{Re}(Px-x,Px+\theta z-Px)=\theta \operatorname{Re}(Px-x,z) \geqslant 0.$$

从而 $\operatorname{Re}(Px-x,z)=0$. 因为对任何 $\theta \in \mathbb{R}, Px+i\theta z \in Z$, 由 (b) 中所得的同一个不等式可得对任何 $\theta \in \mathbb{R}$,

$$\operatorname{Re}(Px-x,Px-i\theta z-Px)=\theta \operatorname{Im}(Px-x,z) \geqslant 0,$$

由此可得 $\operatorname{Im}(Px-x,z)=0$. 因此, 在复的情况下 $(Px-x,z)=0$ 同样成立. 因为

$$(y-x,z-y)=0=\operatorname{Re}(y-x,z-y) \geqslant 0,$$

所以当 $\mathbb{K}=\mathbb{C}$ 时逆命题同样成立. 这就证得 (d).

(v) 先设 Z 为一个子空间. 给定 $x_1,x_2 \in X$ 和 $\alpha_1,\alpha_2 \in \mathbb{K}$. 由 (d) 可知对任何 $z \in Z$.

$$(Px_1-x_1,z)=(Px_2-x_2,z)=0.$$

因此, 对任何 $z \in Z$,

$$((\alpha_1 Px_1+\alpha_2 Px_2)-(\alpha_1 x_1+\alpha_2 x_2),z)=0.$$

由于此时 $\alpha_1 Px_1+\alpha_2 Px_2 \in Z$, 由 (d) 中所建立的特征性质可得

$$\alpha_1 Px_1+\alpha_2 Px_2=P(\alpha_1 x_1+\alpha_2 x_2).$$

所以映射 $P:X \to Z$ 是线性的.

反之, 设 $P:X \to Z$ 是线性的. 因为在映射 P 下 X 的直接像是 Z (显然, 当 $x \in Z$ 时 $Px=x$), 而线性映射的直接像必定是向量空间, 所以集合 Z 是子空间.

最后, 如果 Z 是一个子空间, 在 (c) 的不等式中令 $x_2=0$, 因为 $Px_2=x_2=0 \in Z$, 所以对任何 $x \in X$,

$$\|Px\| \leqslant \|x\|.$$

因而, 除了 $Z=\{0\}$ 的情况, 均有 $\|P\|_{\mathcal{L}(X;Z)}=1$. 这就证得 (e). □

注 如果 Z 是空间 X 的非空子集, 由证明直接可知性质 (b) 和 (d) 的逆命题成立.

后面将对投影定理作若干说明 (也见习题 4.3-1 到 4.3-3).

设 $(X,(\cdot,\cdot))$ 为 Hilbert 空间, 假设 “Z 是完备的” 当然等价于 “Z 在 X 中是闭的”.

在 $X=\mathbb{R}^2,(\cdot,\cdot)$ 是 Euclid 内积的特殊情况下, 定理 4.3-1(a) 中定义的元素 $Px\in Z$ 的几何解释是明显的: Px 是 Z 中离 x “最近” 的元素 (图 4.3-1). 另外, 对任何 $z\in Z$, 两个向量 $Px-x$ 和 $z-Px$ 的夹角的绝对值 $\leqslant\frac{\pi}{2}$ (见定理 4.3-1(b)), 而如果 Z 是一个子空间, 向量 $Px-x$ 直交于任何向量 $z\in Z$, 这又等价于向量 $Px-x$ 与任何向量 $z\in Z$ 的夹角的绝对值等于 $\frac{\pi}{2}$ (见定理 4.3-1 (d)).

正因为如此, 称元素 $Px\in Z$ 为 $x\in X$ 在集合 Z 上的投影, 称算子 $P:X\to Z$ 为 X 到 Z 上的投影算子.

在 (c) 中建立的投影算子 P 带有常数 1 的 Lipschitz 连续性也反映了投影算子另一个显然直观的性质 “投影不增加距离” (图 4.3-1).

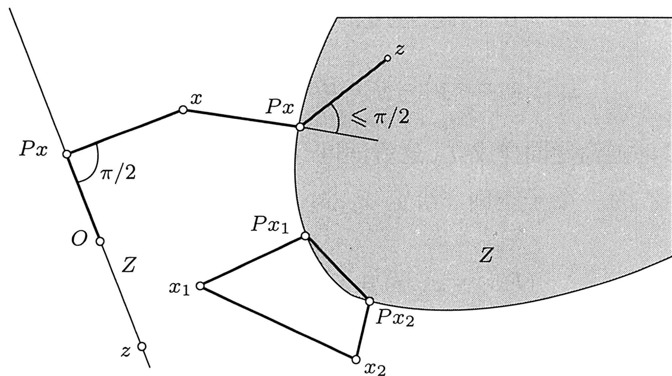


图 4.3-1 当 X 为空间 $\mathbb{R}^2$, 并赋以 Euclid 内积时, 元素 $x\in X$ 的投影 $Px\in Z$ 及定理 4.3-1 所建立的一系列性质的几何解释

必须强调当 Z 是子空间时, 投影算子 $P:X\to Z$ 的线性性质关键在于 (空间 X 中的) 范数是由相应的内积导出的, 见习题 4.3-4.

下面给出投影算子的一些例子. 在赋以定义为 $(\boldsymbol{x},\boldsymbol{y}):=\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$ (这里使用了矩阵的记号, 即将向量视为列向量, 亦即 $n\times 1$ 矩阵) 的 Euclid 内积的空间 $\mathbb{R}^n$ (4.2 节) 中, 考察超平面

$$\boldsymbol{Z}:=\{\boldsymbol{z}\in\mathbb{R}^n;\boldsymbol{a}^{\mathrm{T}}\boldsymbol{z}=0\},$$

即 $\mathbb{R}^n$ 中直交于单位向量 $\boldsymbol{a}\in\mathbb{R}^n(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}=1)$ 的所有向量组成的子空间. 映射

$$\boldsymbol{P}:=\boldsymbol{I}-\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}$$

(这里恒同于一个 $n\times n$ 阵) 是由 $\mathbb{R}^n$ 平行于向量 $\boldsymbol{a}$ 的到超平面 $\boldsymbol{Z}$ 上的投影算

子 (图 4.3-2). 为证明这一点, 注意对所有的 $\boldsymbol{x} \in \mathbb{R}^n$,

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{P}\boldsymbol{x} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} - \boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = 0,$$

且对所有的 $\boldsymbol{z} \in \boldsymbol{Z}$,

$$(\boldsymbol{P}\boldsymbol{x} - \boldsymbol{x})^{\mathrm{T}}\boldsymbol{z} = -\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}\boldsymbol{z} = 0,$$

因此, 对所有的 $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{P}\boldsymbol{x} \in \boldsymbol{Z}$. 由定理 4.3-1(d) 即得所要证明的结论.

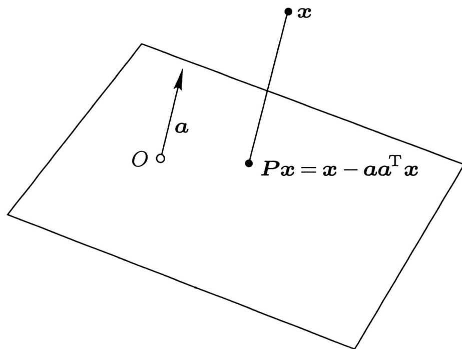


图 4.3-2 投影, 由 $\mathbb{R}^3$ 平行于给定向量 $\boldsymbol{a}$ 到 $\mathbb{R}^3$ 的一个超平面上

这个例子可以直接推广到任意的 Hilbert 空间 $(X, (\cdot,\cdot))$ 和超平面

$$Z := \{z \in X; (a, z) = 0\}$$

上, 其中 a 是 X 中满足 $(a, a) = 1$ 的元素 (集合 Z 显然是 X 的闭子空间). 这里, 平行于 a 的由 X 到 Z 上的投影算子由

$$Px = x - (a, x)a, \quad x \in X$$

给出.

下面考察实 Hilbert 空间 $L^2(\Omega)$, 这里 Ω 是 $\mathbb{R}^n$ 的一个开子集, 令

$$Z := \{g \in L^2(\Omega); \text{在 } A \text{ 上几乎处处 } g = 0\},$$

其中 A 为 Ω 的可测子集. 因为在任何 $L^2(\Omega), 1 \leqslant p < \infty$, 中收敛的任何序列, 均包含几乎处处收敛于同一个极限的子序列 (定理 3.4-3), 所以 Z 是 $L^2(\Omega)$ 的一个子空间, 且在 $L^2(\Omega)$ 中是闭的. 于是, 投影算子 $P : L^2(\Omega) \to Z$ 由

$$Pf = f\chi_{\Omega-A}, \quad f \in L^2(\Omega),$$

给出, 其中 $\chi_{\Omega-A}$ 表示集合 $\Omega - A$ 的特征函数 (图 4.3-3). 为说明这一点, 只要注意到在 $\Omega - A$ 上, 几乎处处 $Pf - f = 0$. 且在 A 上, 几乎处处 $g = 0$, 所以, 对所有

的 $f \in L^2(\Omega), Pf \in Z$, 且对所有的 $g \in Z$,

$$\int_{\Omega}(Pf - f)gdx = 0.$$

因此, 由定理 4.3-1(d) 可知结论成立.

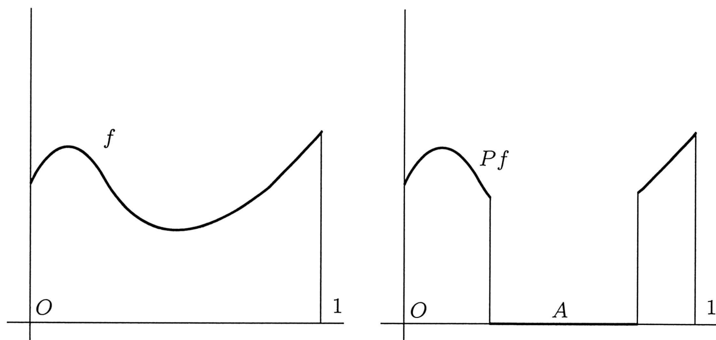


图 4.3-3 当 $\Omega =]0,1[$ 时, 由 $L^2(\Omega)$ 到 $Z = \{g \in L^2(\Omega)$; 在 A 上几乎处处 $g = 0\}$ 上的投影算子

在上述两个例子中, 投影算子均为线性的 (在这两种情况中, Z 均为子空间, 见定理 4.3-1(e)), 下面的例子是非线性的. 如同第一个例子, 空间 X 为赋以 Euclid 内积 $(\cdot,\cdot)$ 的 $\mathbb{R}^n$, 而这里的子集 Z 定义为

$$\mathbb{R}^n_+ := \{(z_i)_{i=1}^n \in \mathbb{R}^n; z_i \geqslant 0, 1 \leqslant i \leqslant n\}.$$

集合 $\mathbb{R}^n_+$ 显然是 $\mathbb{R}^n$ 中的非空凸闭子集, 但并非子空间, 有时也称为 "非负超卦限".

由二维情况 (图 4.3–4) 的启示, 直观上易知对任何向量 $\boldsymbol{x} = (x_i) \in \mathbb{R}^n$, 投影 $\boldsymbol{Px} \in \mathbb{R}^n_+$ 的第 i 个分量 $(\boldsymbol{Px})_i$ 可由

$$(\boldsymbol{Px})_i = \max\{0, x_i\}, \quad 1 \leqslant i \leqslant n,$$

给出.

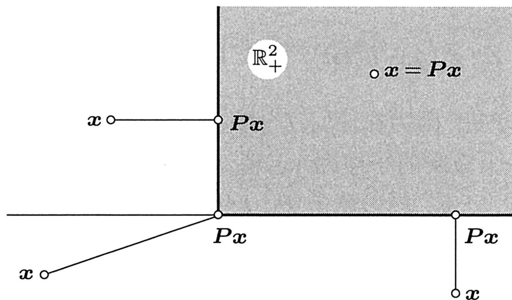


图 4.3-4 由 $\mathbb{R}^2$ 到非负超卦限 $\{(z_i)_{i=1}^2 \in \mathbb{R}^2; z_i \geqslant 0, i = 1, 2\}$ 上的投影. 该图原见于 P.G.CIARLET [2007]: Introduction a l'Analyse Numerique Matricelle et a l'Optimisation. Dunod, Paris.

为验证这个结论, (由定理 2.4–1(b)) 只要证明 $\boldsymbol{Px} \in \mathbb{R}^n_+$, 且对任何 $\boldsymbol{z} \in \mathbb{R}^n_+$, $(\boldsymbol{Px}-\boldsymbol{x}, \boldsymbol{z}-\boldsymbol{Px}) \geqslant 0$. 前者是显然成立的, 后者则因为对任何 $\boldsymbol{z} = (z_i) \in \mathbb{R}^n_+$,

$$(\boldsymbol{Px}-\boldsymbol{x}, \boldsymbol{z}-\boldsymbol{Px}) = \sum_{i=1}^{n}((\boldsymbol{Px})_i - x_i)(z_i - (\boldsymbol{Px})_i) \geqslant 0$$

(如 $x_i \geqslant 0, (\boldsymbol{Px})_i = x_i$; 如 $x_i < 0, (\boldsymbol{Px})_i - x_i = -x_i > 0$, 而 $z_i - (\boldsymbol{Px})_i = z_i \geqslant 0$).

这个例子容易推广到 $\mathbb{R}^n$ 的形如

$$\boldsymbol{Z} := \{(z_i) \in \mathbb{R}^n; a_i \leqslant z_i \leqslant b_i, 1 \leqslant i \leqslant n\}$$

的子集的情况, 对任何向量 $\boldsymbol{x} \in \mathbb{R}^n$, 投影 $\boldsymbol{Px} \in \boldsymbol{Z}$ 的分量由

$$(\boldsymbol{Px})_i = \min\{\max\{x_i, a_i\}, b_i\}, \quad 1 \leqslant i \leqslant n$$

给出, 当 $\boldsymbol{Z}$ 的定义中某些不等式 $a_i \leqslant x_i \leqslant b_i$ 被取代为 $a_i \leqslant x_i$ 或 $x_i \leqslant b_i$, 甚至不再出现时, 也容易做出相应的修正.

由熟知的可逆矩阵的极分解可以给出一个由有限维内积空间到某个非空但却是非凸的闭子集上投影算子的有趣的例子, 见习题 4.3-5.

下面用投影定理的第一个应用结束本节, 这就是 Hilbert 空间中稠密子空间的一个有趣的特征, 即与其所有元素均直交的唯一向量是零向量.

定理 4.3-2 设 $(X, (\cdot,\cdot))$ 为 Hilbert 空间, Y 为 X 的一个子空间. 则

$$\bar{Y} = X$$

的充要条件为对所有的 $y \in Y$, 满足 $(x, y) = 0$ 的唯一元素 $x \in X$ 为 $x = 0$.

证明 假设 $\bar{Y} \neq X$, 任取元素 $\tilde{x} \in (X - \bar{Y})$. 则 $x := \tilde{x} - P\tilde{x}$ 不是零向量, 其中 P 是 X 到 $\bar{Y}$ 上的投影算子; 由投影定理 (定理 4.3-1(d)), $(x, y) = 0$ 对所有的 $y \in \bar{Y}$ 成立, 从而对所有的 $y \in Y$ 成立. 这就证明了充分性部分.

再设 $\bar{Y} = X$, 设向量 $x \in X$ 满足对所有的 $y \in Y, (x, y) = 0$. 因为 $\bar{Y} = X$, 故而存在 $y_n \in Y, n \geqslant 0$, 使得 $\lim_{n\to\infty} y_n = x$. 因此, $(x, x) = \lim_{n\to\infty}(x, y_n) = 0$. 注意这里 "必要性部分" 不依赖 X 的完备性. □

注 实际上, 在任何赋范向量空间上, 用 X' 和 X 间的对偶性取代内积, 类似的性质依然成立 (但此时的证明要用到 Hahn-Banach 定理, 见定理 5.9–4).

习题

4.3-1 设定理 4.3-1 的所有假设中, 除了 Z 非凸以外均满足.

(1) 给出投影唯一性的反例.

(2) 给出投影存在性的反例.

4.3-2　设 X 为 Hilbert 空间, $Z_n, n \geqslant 1$ 为 X 中一列非空的凸闭子集, 满足 $Z_1 \supset Z_2 \supset \cdots \supset Z_n \supset \cdots$. 给定元素 $x \in X$, 用 y_n 表示 x 在 Z_n 上的投影, 其中 $n \geqslant 1$.

(1) 如果 $Z := \cap_{n=1}^{\infty} Z_n \neq \varnothing$, 证明当 $n \to \infty$ 时 $y_n \to y$, 其中 y 是 x 在 Z 上的投影.

(2) 如果 $Z = \varnothing$, 证明当 $n \to \infty$ 时 $\|x - y_n\| \to \infty$.

4.3-3　设 X 为 Hilbert 空间, $Z_n, n \geqslant 1$, 为 X 的一列非空凸闭集, 满足 $Z_1 \subset Z_2 \subset \cdots \subset Z_n \subset \cdots$. 给定元素 $x \in X$, 用 y_n 表示 x 在 Z_n 上的投影, 其中 $n \geqslant 1$.

证明当 $n \to \infty$ 时 $y_n \to y$, 其中 y 表示 x 在集合 $\overline{\cup_{n=1}^{\infty} Z_n}$ 上的投影.

4.3-4　设 $\mathcal{P}_n[0,1] = \{p|_{[0,1]}; p \in \mathcal{P}_n\}$, 其中 $\mathcal{P}_n$ 表示所有次数 $\leqslant n$ 的多项式 $p : \mathbb{R} \to \mathbb{R}$ 组成的空间, 给定数 $q > 1$.

(1) 证明对任意给定的函数 $f \in \mathcal{C}[0,1]$, 存在唯一的多项式 $Pf \in \mathcal{P}_n[0,1]$, 使得

$$\|f - Pf\|_{L^q(0,1)} = \inf_{p \in \mathcal{P}_n[0,1]} \|f - p\|_{L^q(0,1)}.$$

(2) 证明这样定义的映射 $P : \mathcal{C}[0,1] \to \mathcal{P}_n[0,1]$ 是线性的, 当且仅当 $q = 2$ (充分性部分的证明类似于定理 4.3-1(e)).

4.3-5　用 $\mathbb{M}^n, \mathbb{S}^n_>, \mathbb{O}^n$, 和 $\mathbb{O}^n_+$ 分别表示所有 n 阶实的方阵、正定对称阵、正交阵和正常正交阵的集合.

(1) 证明对任意给定的矩阵 $\boldsymbol{A} \in \mathbb{S}^n_>$, 存在唯一的矩阵 $\boldsymbol{B} \in \mathbb{S}^n_>$, 使得 $\boldsymbol{B}^2 = \boldsymbol{A}$. 称矩阵 $\boldsymbol{B}$ 为 $\boldsymbol{A}$ 的平方根, 常记作 $\boldsymbol{A}^{\frac{1}{2}}$.

(2) 证明任何可逆阵 $\boldsymbol{F} \in \mathbb{M}^n$ 可以分解为 $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$. 其中 $\boldsymbol{R} \in \mathbb{Q}^n, \boldsymbol{U} \in \mathbb{S}^n_>$, 且 $\boldsymbol{R}$ 和 $\boldsymbol{U}$ 均是唯一的. 关系式 $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$ 构成了可逆阵 $\boldsymbol{F}$ 的*极分解*.

(3) 设 $\mathbb{U}^n := \{\boldsymbol{F} \in \mathbb{M}^n; \det \boldsymbol{F} \neq 0\}$. 证明按上述定义, 映射 $\boldsymbol{F} \in \mathbb{U}^n \to \boldsymbol{R} \in \mathbb{O}^n$ 和 $\boldsymbol{F} \in \mathbb{U}^n \to \boldsymbol{U} \in \mathbb{S}^n_>$ 均为无穷次可微的 (为此需用到第 7 章中的概念).

(4) 证明 $\mathbb{O}^n_+$ 是 $\mathbb{M}^n$ 中的非空闭子集, 但并非凸的.

(5) 设 $\det \boldsymbol{F} > 0$, 因而 $\boldsymbol{R} \in \mathbb{O}^n_+$. 证明

$$\|\boldsymbol{F} - \boldsymbol{R}\|_F = \inf_{S \in \mathbb{O}^n_+} \|\boldsymbol{F} - \boldsymbol{S}\|_F,$$

其中 $\|\cdot\|_F$ 表示 Frobenius 矩阵范数 (4.2 节).

注　类似于 (2), 可以证明任何可逆的复矩阵可以唯一地分解为一个酉矩阵和一个正定 Hermite 阵的乘积. 此时, 术语 “极分解” 即是复数分解 $z = |z|\mathrm{e}^{i \arg z}$ 的推广.

4.3-6　用 $|\cdot|$ 表示从属于 Euclid 向量范数的矩阵范数 (习题 2.9-1), $\boldsymbol{F} \in \mathbb{M}^n$, 证明

$$\inf_{\boldsymbol{S} \in \mathbb{O}^n_+} |\boldsymbol{F} - \boldsymbol{S}| = |(\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F})^{\frac{1}{2}} - \boldsymbol{I}| \leqslant |\boldsymbol{F}^{\mathrm{T}}\boldsymbol{F} - \boldsymbol{I}|^{\frac{1}{2}}.$$

4.3-7　设 $(X, (\cdot,\cdot))$ 为 Hilbert 空间.

(1) 设 Z 为 X 的闭子空间. 证明相应的连续线性投影算子 $P: X \to Z$ (定理 4.3-1) 具有以下三个性质: $\|P\| = 1$ (除了 $Z = \{0\}$ 的情况), P 是幂等的, 即 $P^2 = P$, 而且 P 是对称的, 即对任何 $x, y \in X, (Px, y) = (x, Py)$.

(2) 设 $Q: X \to X$ 为幂等对称的连续线性算子. 证明 $Q(X)$ 是 X 的闭子空间, 而且 Q 即 X 到 $Q(X)$ 上的投影算子.

(3) 设 $Q: X \to X$ 为连续线性算子, 它是幂等的, 满足 $\|Q\| \leqslant 1$. 证明 $Q(X)$ 是 X 的闭子空间, Q 是 X 到 $Q(X)$ 上的投影算子.

4.3-8 设 X 为 Hilbert 空间, Z 为 X 的闭子空间, $P: X \to Z$ 为相应的投影算子. 证明, 若 $\lambda \neq 0, \lambda \neq 1$, 则连续线性算子 $(\lambda I - P): X \to X$ 为双射, 其逆也是连续的.

4.3-9 设 X 为 Hilbert 空间, 算子 $A \in \mathcal{L}(X)$ 满足 $\|A\| \leqslant 1$. 证明: 对任意的 $x \in X$, 由 $y_n = \frac{1}{n}(x + Ax + \cdots + A^{n-1}x), n \geqslant 1$, 定义的序列 $(y_n)_{n=1}^{\infty}$ 在 X 中收敛.

提示: 证明 $\lim_{n\to\infty} y_n$ 即 x 到 $\operatorname{Span}(A^k x)_{k=0}^{\infty}$ 的闭包上的投影.

4.3-10 设 Y 为实 Hilbert 空间 $(X, (\cdot,\cdot))$ 的非空凸闭子集, $b \in (X - Y)$. 证明存在 $a \in X$ 和 $\alpha \in \mathbb{R}$, 使得对任意的 $y \in X$, 成立

$$(b, y) < \alpha < (y, a).$$

注 这个性质表明超平面 $\{x \in X; (x, a) = \alpha\}$ 严格地分离凸集 Y 和 $\{b\}$. 这一性质实际上在任何赋范向量空间都成立 (但这个一般性结论的证明要复杂些, 见定理 5.10-2).

4.3-11 本题的目标在于建立 Farkaš 引理[6].

设 $(X, (\cdot,\cdot))$ 为实 Hilbert 空间, b 和 $c_i, 1 \leqslant i \leqslant m$ 为 X 中的向量, 则包含关系

$$\{x \in X; (c_i, x) \geqslant 0, 1 \leqslant i \leqslant m\} \subset \{x \in X; (b, x) \geqslant 0\}$$

成立的充要条件是存在实数 $\lambda_i, 1 \leqslant i \leqslant m$, 使得

$$\lambda_i \geqslant 0, \quad 1 \leqslant i \leqslant m \quad 且 \quad b = \sum_{i=1}^{m} \lambda_i c_i.$$

(1) 证明集合

$$Y := \left\{ \sum_{i=1}^{m} \lambda_i c_i \in X; \lambda_i \geqslant 0, 1 \leqslant i \leqslant m \right\}$$

(显然, 它是顶点为 O 的锥) 是 X 的凸闭子集.

(2) 利用问题 (1) 和习题 4.3-10, 证明如果点 $b \in X$ 不属于集合 Y, 则存在向量 $a \in X$, 使得 $(c_i, a) \geqslant 0, 1 \leqslant i \leqslant m$, 且 $(b, a) < 0$.

(3) 利用 (2) 证明 Farkaš 引理的 "必要性" 部分 ("充分性部分" 是显然的).

注 Farkaš 引理在证明 Kuhn-Tucker 乘子的存在性时, 起着关键的作用, 这个乘子出现于当约束为不等式的形式时的约束最优化问题中 (见习题 7.15-3).

---

[6] J. Farkaš [1901]: Theorie der einfachen Ungleichungen. Journal für die Reine und Angewandte Mathematik 124, 1–27.

4.4 投影定理的应用: 线性系统的最小二乘解

对任意给定的 $m \times n$ 实矩阵 $\boldsymbol{A}$ 和任意向量 $\boldsymbol{c} \in \mathbb{R}^m$, 一般并不存在向量 $\boldsymbol{x} \in \mathbb{R}^n$, 使得 $\boldsymbol{Ax} = \boldsymbol{c}$. 取而代之是寻求这个线性系统的最小二乘解[7], 即寻求向量 $\boldsymbol{x} \in \mathbb{R}^n$, 使得 $\mathbb{R}^m$ 中的向量 $\boldsymbol{Ax}$ 和 $\boldsymbol{c}$ 的 Euclid 距离最短 (这就导致术语 "最小二乘" 解). 下面这个投影定理的简单推论显示, 和前一问题不同的是后者至少有一个解.

定理 4.4-1 (线性系统的最小二乘解) 设 $|\cdot|$ 为 $\mathbb{R}^m$ 中的 Euclid 范数.

(a) 给定 $m \times n$ 实矩阵 $\boldsymbol{A}$ 和向量 $\boldsymbol{c} \in \mathbb{R}^m$, 最小化问题, 即求向量 $\boldsymbol{x} \in \mathbb{R}^n$, 使得

$$|\boldsymbol{Ax} - \boldsymbol{c}| = \inf_{\boldsymbol{y} \in \mathbb{R}^n} |\boldsymbol{Ay} - \boldsymbol{c}|,$$

至少有一个解.

(b) 向量 $\boldsymbol{x} \in \mathbb{R}^n$ 满足上述最小化问题当且仅当 $\boldsymbol{x}$ 是下述线性系统

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{c}$$

的解.

证明 因为 $\mathrm{Im}\,\boldsymbol{A}$ 是 $\mathbb{R}^m$ 的闭子空间 (它是有限维子空间, 见定理 2.7-1(c)), 由投影定理 (定理 4.3-1), 存在唯一的向量 $\tilde{\boldsymbol{x}} \in \mathrm{Im}\,\boldsymbol{A}$, 满足

$$|\tilde{\boldsymbol{x}} - \boldsymbol{c}| = \inf_{\tilde{\boldsymbol{y}} \in \mathrm{Im}\,\boldsymbol{A}} |\tilde{\boldsymbol{y}} - \boldsymbol{c}|,$$

这又等价于对任意的 $\tilde{\boldsymbol{y}} \in \mathrm{Im}\,\boldsymbol{A}$,

$$(\tilde{\boldsymbol{x}} - \boldsymbol{c}, \tilde{y})_m = 0,$$

其中 $(\cdot,\cdot)_m$ 表示 $\mathbb{R}^m$ 中的 Euclid 内积. 由空间 $\mathrm{Im}\,\boldsymbol{A}$ 的定义, 至少存在一个向量 $\boldsymbol{x} \in \mathbb{R}^n$, 满足

$$|\boldsymbol{Ax} - \boldsymbol{c}| = \inf_{\boldsymbol{y} \in \mathbb{R}^n} |\boldsymbol{Ay} - \boldsymbol{c}|,$$

这又等价于对任意的 $\boldsymbol{y} \in \mathbb{R}^n$,

$$(\boldsymbol{Ax} - \boldsymbol{c}, \boldsymbol{Ay})_m = (\boldsymbol{A}^{\mathrm{T}}\boldsymbol{Ax} - \boldsymbol{A}^{\mathrm{T}}\boldsymbol{c}, \boldsymbol{y})_n = 0,$$

其中 $(\cdot,\cdot)_n$ 表示 $\mathbb{R}^n$ 的 Euclid 内积, 矩阵 $\boldsymbol{A}^{\mathrm{T}}$ 为 $\boldsymbol{A}$ 的转置阵.

这就证得了 (a) 和 (b). □

---

[7] 这个方法当初是计算 (当然是手算) 天体运动的轨迹发明的, 见:

A. M. Legendre [1805]: Nouvelle Méthode pour la Détermination des Orbites des Comètes. Chez Didot, Paris.

C. F. Gauss [1809]: Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium. Perthes und Besser, Hamburg.

由投影定理, 至少有一个解的线性系统 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{c}$ 构成了与线性系统 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$ 的最小二乘解相关的*正规方程组*[8]. 当然, 如果集合 $\{\boldsymbol{x} \in \mathbb{R}^n; \boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}\}$ 非空, 则它与正规方程组的解集一致.

要强调的是寻求线性系统的最小二乘解之所以导出了一个线性问题 (即正规方程组) 只是因为这里的范数是由内积导出的 (在 4.3 节末尾, 关于到子空间上的投影算子的线性性质有类似的讨论). 由此说明了从数值计算的观点来看, 相比于 "最小 $\|\cdot\|_p$ 范数解", 其中 $p \neq 2$, 人们更偏爱最小二乘解的原因.

注 (1) 上述讨论可以直接推广到复的情况, 此时正规方程组变为 $\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^*\boldsymbol{c}$, 其中 $\boldsymbol{A}^*$ 为 $\boldsymbol{A}$ 的转置伴随矩阵.

(2) 正规方程组解的唯一性的条件在习题 4.7-2 中讨论.

4.5 直交性; 直和定理

设 $(X, (\cdot,\cdot))$ 为实或复的内积空间. 两个向量 $x \in X$ 和 $y \in X$ 称之为*直交的*, 是指它们满足

$$(x, y) = 0.$$

X 的任何非空子集 Z 的*直交补*定义为 X 的子集

$$Z^{\perp} := \{x \in X;\ \text{对所有的}\ z \in Z, (x, z) = 0\}.$$

下面的定理列出了直交补的若干基本性质.

定理 4.5-1 *设 Z 为内积空间 X 的非空子集. 则集合 $Z^{\perp}$ 是 X 的闭子空间. 而且 $(\overline{Z})^{\perp} = Z^{\perp}$, 如果 $0 \in Z$, 则 $Z \cap Z^{\perp} = \{0\}$, 如果 $0 \bar{\in} Z$, 则 $Z \cap Z^{\perp} = \varnothing$.*

证明 由内积关于其第一个变元的线性可知 $Z^{\perp}$ 是一个子空间; 又由内积关于第一个变元的连续性 (定理 4.1-1(c)) 可知 $Z^{\perp}$ 是闭的.

由直交补的定义即得 $(\overline{Z})^{\perp} \subset Z^{\perp}$. 为证明 $Z^{\perp} \subset (\overline{Z})^{\perp}$, 设 $x \in Z^{\perp}$, 由于对所有的 $z \in Z, (x, z) = 0$, 利用内积关于其第二个变元的连续性可得对所有的 $z \in \overline{Z}, (x, z) = 0$, 所以 $x \in (\overline{Z})^{\perp}$.

当 $0 \in Z$ 时, 显然有 $Z \cap Z^{\perp} = \{0\}$; 当 $0 \bar{\in} Z$ 时 $Z \cap Z^{\perp} = \varnothing$ 也是显然的. □

当 X 为 Hilbert 空间, Y 为 X 的闭子空间时, X 可以表示为子空间 Y 和 $Y^{\perp}$ 的*直和* (2.1 节), 后者也是 X 的闭子空间 (定理 4.5-1). 下面可以看到, 这个重要的性质实际上是投影定理的简单推论.

---

[8] 提出并这样命名见:

C. F. Gauss [1822]: Anwendung der Wahrscheinlichkeitsrechnung auf eine Aufgabe der practischen Geometrie. Astronomische Nachrichten 1, 81–86.

定理 4.5-2 (直和定理) 设 X 为实或复的 Hilbert 空间, Y 为 X 的闭子空间, 则 X 可表示为直和

$$X = Y \oplus Y^{\perp},$$

即任何 $x \in X$ 可表示为

$$x = y + y^{\perp},$$

其中 $y \in Y, y^{\perp} \in Y^{\perp}$, 而且这样的分解是唯一的. 实际上

$$y = Px, \quad y^{\perp} = P^{\perp}x,$$

其中 $P: X \to Y$ 表示 X 到 Y 上的投影算子,

$$P^{\perp} := I - P$$

为 X 到 $Y^{\perp}$ 上的投影算子.

证明 任何元素 $x \in X$ 可表示为

$$x = Px + (I - P)x.$$

由投影算子的定义, $Px \in Y$. 由到闭子空间上的投影的特征 (定理 4.3-1(d)), 对任何 $z \in Y, ((I-P)x, z) = 0$, 因此 $(I-P)x \in Y^{\perp}$. 这样,

$$x = y + y^{\perp},$$

其中

$$y := Px \in Y, \quad y^{\perp} := (I-P)x \in Y^{\perp}.$$

为验证分解的唯一性, 设

$$x = y + y^{\perp} = \hat{y} + \hat{y}^{\perp},$$

其中

$$y, \hat{y} \in Y, \quad y^{\perp}, \hat{y}^{\perp} \in Y^{\perp}.$$

因为 $(y - \hat{y}) \in Y, (y^{\perp} - \hat{y}^{\perp}) \in Y^{\perp}$ ($Y^{\perp}$ 也是子空间, 见定理 4.5-1), 又因 $Y \cap Y^{\perp} = \{0\}$, 所以 $y - \hat{y} = y^{\perp} - \hat{y}^{\perp} = 0$.

由投影算子的特征, 对任何 $x \in X$, 因 $Px \in Y$, 故而对任何 $y^{\perp} \in Y^{\perp}$, 有

$$(x - P^{\perp}x, y^{\perp}) = (Px, y^{\perp}) = 0,$$

所以 $P^{\perp} := I - P$ 是 X 到子空间 $Y^{\perp}$ 上的投影算子. □

注 (1) 如果 Y 是闭子空间, 由定理 4.5-2, 因为 $X = Y \oplus Y^{\perp} = (Y^{\perp})^{\perp} \oplus Y^{\perp}$, 所以 $Y = (Y^{\perp})^{\perp}$.

(2) 如果 Y 为未必是闭的子空间, 则因 $(\bar{Y})^{\perp} = Y^{\perp}$ (定理 4.5-1), X 仍可表示为直和 $X = \bar{Y} \oplus Y^{\perp}$.

习题

4.5-1 在空间 $\mathcal{C}^1[0,1]$ 上赋以定义为

$$(f,g) := \int_0^1 (f'g' + fg)dx$$

的内积 $(\cdot,\cdot)$, 在 $\mathcal{C}^1[0,1]$ 上定义子集 Y 为

$$Y := \{g \in \mathcal{C}^1[0,1]; g(0) = g(1) = 0\}.$$

(1) 证明 Y 是 $(\mathcal{C}^1[0,1],(\cdot,\cdot))$ 也是 $(\mathcal{C}^2[0,1],(\cdot,\cdot))$ 的闭子空间.

提示: 证明存在常数 C 使得对所有 $f \in \mathcal{C}^1[0,1]$ 有 $\sup\limits_{0\leqslant x\leqslant 1} |f(x)| \leqslant C(f,f)^{\frac{1}{2}}$.

(2) 在 $(\mathcal{C}^2[0,1],(\cdot,\cdot))$ 中求 Y 的直交补 $Y^\perp$. $Y^\perp$ 的维数是多少?

4.5-2 设 Hilbert 空间 ℓ^2 (4.2 节) 的子集 Y 定义为

$$Y := \{x = (x_i)_{i=1}^\infty;\ \text{对所有的整数}\ k \geqslant 1, x_{2k-1} = x_{2k}\}.$$

(1) 证明 Y 是 ℓ^2 的闭子空间.

(2) 求出 Y 在 ℓ^2 中的直交补.

(3) 求出投影算子 $P : \ell^2 \to Y$ 和 $P^\perp : \ell^2 \to Y^\perp$.

4.6 Hilbert 空间中的 F. Riesz 表示定理

设 $(X,(\cdot,\cdot))$ 为 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的内积空间, 用 X' 表示其对偶空间. 任意给定向量 $y \in X$, 定义线性泛函 $\ell_y : X \to \mathbb{K}$ 为

$$\ell_y(x) := (x,y) \in \mathbb{K}, \quad \forall x \in X,$$

它是连续的, 而且

$$\|\ell_y\|_{X'} = \|y\|,$$

这是因为如果 $y \neq 0$, 则由定理 4.1-1 可得

$$\|\ell_y\|_{X'} = \sup_{x\neq 0} \frac{|\ell_y(x)|}{\|x\|} = \sup_{x\neq 0} \frac{|(x,y)|}{\|x\|} = \|y\|.$$

值得注意的十分重要的一点是, 由直和定理 (其本身是投影定理的推论), 当 X 为 Hilbert 空间时, 上述事实的逆也成立.

定理 4.6-1 (Hilbert 空间中的 F. Riesz 表示定理[9]) 设 $(X,(\cdot,\cdot))$ 为 $\mathbb{K} = \mathbb{R}$ 或 $\mathbb{K} = \mathbb{C}$ 上的 Hilbert 空间. 对任意给定的连续线性泛函 $\ell \in X^*$, 存在唯一的向

[9] F. Riesz [1907]: Sur une espèce de géométrie analytique des systèmes de fonctions sommables. Comptes Rendus de l'Académie des Sciences de Paris 144, 1409–1411.

量 $y_\ell \in X$, 使得对所有的 $x \in X$,

$$\ell(x) = (x, y_\ell).$$

而且

$$\|\ell\|_{X'} = \|y_\ell\|_X,$$

由此定义的 F. Riesz 等距算子

$$\sigma : \ell \in X' \to \sigma(\ell) = y_\ell \in X$$

是一个双射, 当 $\mathbb{K} = \mathbb{R}$ 时, 它是线性的; 当 $\mathbb{K} = \mathbb{C}$ 时, 它是半线性的.

由此, 借助于 F.Riesz 等距算子 $\sigma : X' \to X$, 任何 Hilbert 空间 X 均可恒同于其对偶空间; 而且, 在对偶空间 X' 上赋以内积 $(\cdot,\cdot)_{X'} : X' \times X' \to \mathbb{K}$, 其定义为对任意 $x', y' \in X'$,

$$(x', y')_{X'} := \overline{(\sigma x', \sigma y')},$$

则 X' 也是 Hilbert 空间.

证明 如果 $\ell = 0$, 取 $y_\ell = 0$ 即可. 如果 $\ell \neq 0$, 令

$$Y := \{x \in X; \ell(x) = 0\}.$$

由 $\ell : X \to \mathbb{K}$ 的连续性可知 Y 为一闭子空间, 又由 $\ell \neq 0$ 可知 $Y \subsetneqq X$. 因此, 由直积定理 (定理 4.5-2),

$$X = Y \oplus Y^\perp,$$

且 $Y^\perp$ 包含非零向量 ($Y^\perp = \{0\}$ 将导致 $Y = X$). 取 $y_0 \in Y^\perp$ 满足 $y_0 \neq 0$, 于是 $\ell(y_0) \neq 0$ (否则, 如 $\ell(y_0) = 0$, 则 $y_0 \in Y$, 但 $Y \cap Y^\perp = \{0\}$). 因而不失一般性, 假设

$$\ell(y_0) = 1.$$

由到闭子空间上的投影的特征 (定理 4.3-1(d)), 投影算子 $P : X \to Y$ 可表示为对任意的 $x \in X$,

$$Px = x - \ell(x) y_0,$$

这是因为 $Px \in Y$, 且对任意的 $y \in Y, (Px - x, y) = -\ell(x)(y_0, y) = 0$.

由此, 投影算子 $P^\perp : X \to Y^\perp$ 可表示为 (定理 4.5-2) 对任意的 $x \in X$,

$$P^\perp x = (I - P)x = \ell(x) y_0.$$

取向量

$$y_\ell := \frac{1}{\|y_0\|^2} y_0.$$

则对所有的 $x \in X$,

$$(x, y_\ell) = \frac{1}{\|y_0\|^2}(Px + P^\perp x, y_0) = \frac{1}{\|y_0\|^2}(P^\perp x, y_0) = \ell(x),$$

即它满足定理所述的性质.

因为对所有的 $x \in X$ 均有当 $(x, y) = (x, \tilde{y})$ 时, 必有 $y = \tilde{y}$ (取 $x = y - \tilde{y}$ 即可), 所以上述的 y_ℓ 是唯一确定的. 显然, 这样定义的映射 $\ell \in X' \to y_\ell \in X$ 是双射, 当 $\mathbb{K} = \mathbb{R}$ 时它是线性的, 当 $\mathbb{K} = \mathbb{C}$ 时, 它是半线性的. 另一方面

$$\|\ell\|_{X'} = \sup_{x \neq 0} \frac{|\ell(x)|}{\|x\|} = \sup_{x \neq 0} \frac{|(x, y_\ell)|}{\|x\|} = \|y_\ell\|.$$

最后, 作为 X 上内积 $(\cdot,\cdot)$ 半双线性的推论, 可以直接验证函数 $(\cdot,\cdot)_{X'} : X' \times X' \to \mathbb{K}$ (如定理叙述所定义的) 是 X' 上的内积, 因为对任何 $x' \in X'$,

$$\|x'\|_{X'} = (x', x')_{X'}^{\frac{1}{2}} = (\sigma x', \sigma x')^{\frac{1}{2}} = \|\sigma x'\|,$$

易知内积空间 $(X', (\cdot,\cdot)_{X'})$ 是完备的. □

注 由在上述证明中建立的关系: 对所有的 $x \in X, P^\perp x = \ell(x) y_0$ 可见

$$Y^\perp = P^\perp(X) = \{\alpha y_0 \in X; \alpha \in \mathbb{K}\} = \mathrm{Span}(y_0)$$

是空间 X 的一维子空间.

例如, 设 Ω 为 $\mathbb{R}^N$ 的开子集, A 为 Ω 的可测子集, 满足 $\int_A dx < \infty$. 于是集合 A 的特征函数 χ_A 属于 (实) Hilbert 空间 $L^2(\Omega)$. 由 Cauchy-Schwarz-Bunyakovskiĭ 不等式, 泛函 $\ell : f \in L^2(\Omega) \to \int_A f(x)dx$ 显然是连续的, 它也可表示为 $f \in L^2(\Omega) \to \int_\Omega \chi_A(x) f(x) dx$.

更一般地, 由定理 4.6-1 可知, 给定空间 $L^2(\Omega)$ 上的任何连续线性泛函 ℓ, 存在函数 $g_\ell \in L^2(\Omega)$, 使得对任何 $f \in L^2(\Omega)$,

$$\ell(f) = \int_\Omega f(x) g_\ell(x) dx.$$

这个值得注意的结果只是 F. Riesz 表示定理的一个不费力气的应用, 其原因在于 $L^2(\Omega)$ 是 Hilbert 空间. 要把这个结果推广到空间 $L^p(\Omega)$, 其中 $1 < p < \infty, p \neq 2$, 则需要更为特殊的复杂而精细的证明 (此时, 函数 g_ℓ 属于 $L^q(\Omega)$, 其中 $\frac{1}{p} + \frac{1}{q} = 1$, 见定理 3.5-3).

4.7 F. Riesz 表示定理的应用: Hilbert 空间中的 Hahn-Banach 定理; 伴随算子; 再生核

F. Riesz 表示定理和直和定理一起提供了线性泛函分析中最基本的结果之一的 "Hilbert 空间形式" 的十分简单的证明, 对于任意的赋范向量空间的情况, 需要用到 "选择公理" (定理 5.9-1). 如前所述, 用 X' 表示赋范向量空间 X 的对偶空间.

定理 4.7-1 (Hilbert 空间中的 Hahn-Banach 定理) 设 X 为 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$ 上的 Hilbert 空间, Y 为 X 的子空间, $\ell: Y \to \mathbb{K}$ 为 Y 上的连续线性型, 则存在连续线性型 $\tilde{\ell}: X \to \mathbb{K}$, 满足对所有的 $y \in Y$,

$$\tilde{\ell}(y)=\ell(y) \quad 且 \quad \|\tilde{\ell}\|_{X'}=\|\ell\|_{Y'},$$

而且, 这样的延拓是唯一的.

证明 用 $\bar{Y}$ 表示 Y 在 X 中的闭包. 因为域 $\mathbb{K}$ 是完备的. 所以存在唯一的连续线性型 $\hat{\ell}: \bar{y} \to \mathbb{K}$, 满足对所有的 $y \in Y$.

$$\hat{\ell}(y)=\ell(y), \quad 而且 \quad \|\hat{\ell}\|_{(\bar{Y})'}=\|\ell\|_{Y'}$$

(定理 3.1-1). 由直和定理 (定理 4.5-2), 任何元素 $x \in X$ 可唯一地表示为

$$x=Px+P^{\perp}x,$$

这里的 P 和 $P^{\perp}$ 分别表示 Hilbert 空间 X 到其闭子空间 $\bar{Y}$ 和 $(\bar{Y})^{\perp}$ 上的投影算子. 定义线性型 $\tilde{\ell}: X \to \mathbb{K}$ 为

$$\tilde{\ell}(x):=\hat{\ell}(Px), \quad x \in X.$$

因为对所有的 $y \in Y$,

$$\tilde{\ell}(y)=\hat{\ell}(Py)=\hat{\ell}(y)=\ell(y),$$

所以 $\tilde{\ell}$ 是 ℓ 的延拓. 又因为对所有的 $x \in X, \|Px\| \leqslant \|x\|$,

$$\begin{aligned}\|\ell\|_{Y'} &= \sup_{\substack{y \in Y \\ y \neq 0}} \frac{|\ell(y)|}{\|y\|} \leqslant \sup_{\substack{x \in X \\ x \neq 0}} \frac{|\tilde{\ell}(x)|}{\|x\|}=\|\tilde{\ell}\|_{X'} \\ &= \sup_{\substack{x \in X \\ x \neq 0}} \frac{|\hat{\ell}(Px)|}{\|x\|} \leqslant \|\hat{\ell}\|_{\overline{Y}'}=\|\ell\|_{Y'},\end{aligned}$$

所以 $\|\tilde{\ell}\|_{X'}=\|\ell\|_{Y'}$.

为验证这样的延拓是唯一的, 不失一般性, 不妨设 Y 是闭的 (因为由 Y 到 $\bar{Y}$ 上的延拓是唯一的). 现在设 $\ell^{\#} \in X'$ 是 $\ell \in Y'$ 的一个延拓, 满足 $\|\ell^{\#}\|_{X'}=\|\ell\|_{Y'}$. 于是, 由 Hilbert 空间的 F. Riesz 表示定理, 存在唯一的向量 $z \in X$, 满足对所有的 $x \in X$.

$$\ell^{\#}(x)=(x, z) \quad 且 \quad \|\ell^{\#}\|_{X'}=\|z\|.$$

因为对所有的 $y \in Y$,

$$\ell^{\#}(y) = \ell(y) = (y, z) = (y, Pz),$$

所以 $\|\ell\|_{Y'} = \|Pz\|$. 从而由 $\|\ell^{\#}\|_{X'} = \|\ell\|_{Y'}$ 得 $\|Pz\| = \|z\|$, 因此 $z = Pz$. 于是, 对所有的 $x \in X$,

$$\ell^{\#}(x) = (x, Pz) = (Px, z) = \ell^{\#}(Px) = \ell(Px),$$

由此即得 $\ell^{\#} = \tilde{\ell}$. □

作为对 Hilbert 空间上 F. Riesz 表示定理另一个应用的准备, 考察分别赋以 Euclid 内积 $(\cdot,\cdot)_n$ 和 $(\cdot,\cdot)_m$ 的空间 $\mathbb{R}^n$ 和 $\mathbb{R}^m$ (4.2 节). 对任何实 $m \times n$ 阵 $\boldsymbol{A} = (a_{ij})$, 其 $n \times m$ 转置阵 $\boldsymbol{A}^{\mathrm{T}}$ 定义为 $(\boldsymbol{A}^{\mathrm{T}})_{ij} = a_{ji}$, 也可定义为唯一的对所有的 $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m$ 成立

$$(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{y})_m = (\boldsymbol{x}, \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y})_n$$

的 $n \times m$ 阵. 类似地, 对任何复 $m \times n$ 阵 $\boldsymbol{A} = (a_{ij})$, 其 $n \times m$ 伴随矩阵 $\boldsymbol{A}^*$ 也可定义为唯一的对所有的 $\boldsymbol{x} \in \mathbb{C}^n, \boldsymbol{y} \in \mathbb{C}^m$ 成立

$$(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{y})_m = (\boldsymbol{x}, \boldsymbol{A}^*\boldsymbol{y})_n$$

的 $n \times m$ 复矩阵, 其中 $(\cdot,\cdot)_n$ 和 $(\cdot,\cdot)_m$ 分别表示 $\mathbb{C}^n$ 和 $\mathbb{C}^m$ 上的 Hermite 内积 (4.2 节).

值得注意是由于 F. Riesz 表示定理, 对于两个 Hilbert 空间之间的连续性算子, 可以类似于上述实空间下的转置或复空间下的伴随, 导出相应的概念, 为简明起见, 下面的定理只讨论复空间的情况, 在证明后将对实空间情况下的修正作一些说明. 若干补充将在习题 4.7-1 中给出.

定理 4.7-2 (伴随算子) 设 $(X, (\cdot,\cdot)_X)$ 和 $(Y, (\cdot,\cdot)_Y)$ 为两个复 Hilbert 空间, 给定算子 $A \in \mathcal{L}(X; Y)$.

(a) 存在唯一的算子 $A^* \in \mathcal{L}(Y; X)$, 称之为 A 的伴随算子, 对所有的 $x \in X, y \in Y$, 满足

$$(Ax, y)_Y = (x, A^*y)_X.$$

这样定义的映射 $A \in \mathcal{L}(X; Y) \to A^* \in \mathcal{L}(Y; X)$ 是半线性的. 而且

$$\|A^*\|_{\mathcal{L}(Y;X)} = \|A\|_{\mathcal{L}(X;Y)}.$$

(b) 下列关系式成立:

$$(\mathrm{Im}\ A)^{\perp} = \mathrm{Ker}\ A^*, \quad (\mathrm{Im}\ A^*)^{\perp} = \mathrm{Ker}\ A,$$

$$Y = \mathrm{Ker}\ A^* \oplus \overline{\mathrm{Im}\ A}, \quad X = \mathrm{Ker}\ A \oplus \overline{\mathrm{Im}\ A^*}.$$

证明 (a) 对每个 $y \in Y$, 因为 $|(Ax,y)_Y| \leqslant \|A\|\|x\|\|y\|$ 对所有的 $x \in X$ 成立, 所以映射 $x \in X \to (Ax,y)_Y \in \mathbb{K}$ 是连续线性泛函. 在 Hilbert 空间 X 上用 F. Riesz 表示定理 (定理 4.6-1), 即得存在唯一的向量 $A^*y \in X$, 使得对一切 $x \in X$, 成立

$$(Ax,y)_Y = (x, A^*y)_X.$$

这样定义的映射 $A^* : Y \to X$ 是线性的, 这是因为对任何 $\alpha, \beta \in \mathbb{C}, x \in X$ 和 $y, z \in Y$,

$$\begin{aligned}(x, A^*(\alpha y + \beta z)) &= (Ax, \alpha y + \beta z) = \bar{\alpha}(Ax,y) + \bar{\beta}(Ax,z) \\ &= \bar{\alpha}(x, A^*y) + \bar{\beta}(x, A^*z) = (x, \alpha A^*y + \beta A^*z).\end{aligned}$$

对任何 $A, B \in \mathcal{L}(X;Y)$, 显然有 $(\alpha A + \beta B)^* = \bar{\alpha}A^* + \bar{\beta}B^*$.

线性算子 $A^* : Y \to X$ 是连续的. 这是因为对任何 $y \in Y$,

$$\|A^*y\|^2 = (A^*y, A^*y)_X = (AA^*y, y)_Y \leqslant \|A\|\|A^*y\|\|y\|,$$

从而

$$\|A^*\|_{\mathcal{L}(Y;X)} = \sup_{y \neq 0} \frac{\|A^*y\|}{\|y\|} \leqslant \|A\|_{\mathcal{L}(X;Y)}.$$

类似地, 对所有的 $x \in X$,

$$\|Ax\|^2 = (Ax, Ax)_Y = (x, A^*Ax)_X \leqslant \|A^*\|\|Ax\|\|x\|,$$

从而

$$\|A\|_{\mathcal{L}(X;Y)} = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \leqslant \|A^*\|_{\mathcal{L}(Y;X)}.$$

因此, $\|A^*\|_{\mathcal{L}(Y;X)} = \|A\|_{\mathcal{L}(X;Y)}$. 这就证得了 (a).

再证 (b), 注意

$$\begin{aligned}(\mathrm{Im}\ A)^{\perp} &= \{y \in Y;\ \text{对所有的 } z \in \mathrm{Im}\ A, (y,z)_Y = 0\} \\ &= \{y \in Y;\ \text{对所有的 } x \in X, (y, Ax)_Y = 0\} \\ &= \{y \in Y;\ \text{对所有的 } x \in X, (A^*y, x)_X = 0\} = \mathrm{Ker}\ A^*.\end{aligned}$$

因为 $(\overline{\mathrm{Im}\ A})^{\perp} = \mathrm{Im}\ A^{\perp}$ (定理 4.5-1), 由直和定理 (定理 4.5-2) 可得

$$Y = \overline{\mathrm{Im}\ A} \oplus (\overline{\mathrm{Im}\ A})^{\perp} = \overline{\mathrm{Im}\ A} \oplus (\mathrm{Im}\ A)^{\perp} = \overline{\mathrm{Im}\ A} \oplus \mathrm{Ker}\ A^*.$$

类似可得 (b) 中的另一关系式. □

注 (1) 为建立伴随算子 A^* 的存在性, 并不需要空间 Y 的完备性. 只要 Y 可以表示为直和 $Y = \mathrm{Ker}\ A^* \oplus \overline{\mathrm{Im}\ A}$ 就能导出结论.

(2) 自然地, 若 Y 是有限维的, 则 $\overline{\mathrm{Im}\ A} = \mathrm{Im}\ A$, 若 X 是有限维的, 则 $\overline{\mathrm{Im}\ A^*} = \mathrm{Im}\ A^*$.

如果 $(X,(\cdot,\cdot)_X)$ 和 $(Y,(\cdot,\cdot)_Y)$ 为实 Hilbert 空间, 类似地可证明对任意给定的算子 $A \in \mathcal{L}(X;Y)$, 存在唯一的算子 $A^{\mathrm{T}} \in \mathcal{L}(Y;X)$, 称之为 A 的转置, 使得对所有的 $x \in X, y \in Y$,

$$(Ax,y)_Y = (x, A^{\mathrm{T}}y)_X.$$

除了这样定义的映射 $A \in \mathcal{L}(X;Y) \to A^{\mathrm{T}} \in \mathcal{L}(Y;X)$ 是线性的而外, 定理 4.7-2 中建立的其他性质在用 A^{T} 取代 A^* 后依然成立.

上述定理的一个简单推论是矩阵论中下面的经典结果.

定理 4.7-3 (有限维空间中的 Fredholm 两择性) 假设给定实 $(\mathbb{K}=\mathbb{R})$ 或复 $(\mathbb{K}=\mathbb{C})$ 的 $m\times n$ 阵 $\boldsymbol{A}$ 和向量 $\boldsymbol{b} \in \mathbb{K}^m$.

那么或者线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 至少有一个解 $\boldsymbol{x} \in \mathbb{K}^n$; 或者方程组无解, 此时, 至少存在一个向量 $\boldsymbol{y} \in \mathbb{K}^m$, 使得当 $\mathbb{K}=\mathbb{R}$ 时,

$$\boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = 0 \quad 且 \quad \boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} \neq 0.$$

或当 $\mathbb{K}=\mathbb{C}$ 时,

$$\boldsymbol{A}^*\boldsymbol{y} = 0 \quad 且 \quad \boldsymbol{y}^*\boldsymbol{b} \neq 0.$$

证明 为确定起见, 设 $\mathbb{K}=\mathbb{C}$ (当 $\mathbb{K}=\mathbb{R}$ 时证明是类似的), 在 $\mathbb{C}^n$ 上赋以 Hermite 内积 (4.2 节). 注意, 有限维空间 $\mathrm{Im}\,A$ 是闭的 (定理 2.7-1(c)), 由定理 4.7-2(b) 可得

$$\mathbb{C}^m = \mathrm{Ker}\,\boldsymbol{A}^* \oplus \mathrm{Im}\,\boldsymbol{A}.$$

于是, 或者 $\boldsymbol{b} \in \mathrm{Im}\,\boldsymbol{A}$, 此时线性方程组 $\boldsymbol{Ax}=\boldsymbol{b}$ 至少有一个解, 或者 $\boldsymbol{b} \bar{\in} \mathrm{Im}\,\boldsymbol{A}$, 此时线性方程组无解. 又由 $\boldsymbol{b} \bar{\in} \mathrm{Im}\,A$ 可知 $\boldsymbol{b}$ 在空间 $\mathrm{Ker}\,A^*$ 上的投影 $\boldsymbol{y}$ 不是 $\mathbb{C}^m$ 中的零向量, 它满足 $\boldsymbol{A}^*\boldsymbol{y}=\boldsymbol{0}$, 且 $\boldsymbol{y}^*\boldsymbol{b}=\boldsymbol{y}^*\boldsymbol{y}\neq 0$. □

作为另一个应用的准备, 考察空间 ℓ^2, 其元素 $x=(x_i)_{i=0}^{\infty}$ 实际上即函数 $x: i \in \mathbb{N} \to \mathbb{K}$. 对每个 $j \in \mathbb{N}$, 令 $e_j = (\delta_{ij})_{i=0}^{\infty}$ 显然, 对每个 $j \in \mathbb{N}, e_j \in \ell^2$, 且对所有的 $j \geqslant 0$ 的所有 $x=(x_i) \in \ell^2$,

$$x_j = (x, e_j)_{\ell^2}.$$

F. Riesz 表示定理另一个推论提供了一条简单的准则以保证元素为函数的更一般的 Hilbert 空间也具有上述性质.

定理 4.7-4 (再生核) 设 A 为非空集合, $(X,(\cdot,\cdot))$ 为 $\mathbb{K}=\mathbb{R}$ 或 $\mathbb{K}=\mathbb{C}$ 上的 Hilbert 空间, 其元素为函数 $x: A \to \mathbb{K}$. 设对每个 $a \in A$, 存在数 $C(a)>0$, 使得对任何 $x \in X$,

$$|x(a)| \leqslant C(a)\|x\|.$$

则存在函数

$$K: A \times A \to \mathbb{K},$$

称之为 X 的再生核, 使得对每个 $a \in A$, 函数 $K(\cdot, a): A \to \mathbb{K}$ 是空间 X 的元素, 对任何 $x \in X$,

$$x(a) = (x, K(\cdot, a)).$$

证明 由假设, 对每个 $a \in A$, 线性泛函 $x \in X \to x(a) \in \mathbb{K}$ 是连续的. 由 F. Riesz 表示定理可得存在空间 X 中的元素 $K(\cdot, a)$, 该元素为函数 $K(\cdot, a): A \to \mathbb{K}$, 满足对所有的 $x \in X, x(a) = (x, K(\cdot, a))$. □

F. Riesz 表示定理的这个看似无足轻重的推论实际上有着重要的应用. 特别如用于关于某类边值问题非负 Green 函数存在性的讨论[10).

F. Riesz 表示定理的另一个重要的应用见习题 4.7-3.

习题

4.7-1 记号和假设同定理 4.7-2.

(1) 证明 $(A^*)^* = A$.

(2) 证明 $\mathrm{Ker}\ A^* = \mathrm{Ker}\ (AA^*)$, 且 $\overline{\mathrm{Im}\ A} = \overline{\mathrm{Im}(AA^*)}$.

(3) 证明 $\|A^*A\|_{\mathcal{L}(X)} = \|A\|^2_{\mathcal{L}(X;Y)} = \|AA^*\|_{\mathcal{L}(Y)}$.

(4) 证明: 如果 $A \in \mathcal{L}(X;Y)$ 是双射, 且 $A^{-1} \in \mathcal{L}(Y;X)$, 则 $A^* \in \mathcal{L}(Y;X)$ 是双射, 且 $(A^*)^{-1} \in \mathcal{L}(X;Y)$; 又 $(A^*)^{-1} = (A^{-1})^*$.

(5) 设 $(Z, (\cdot,\cdot)_z)$ 是另一个复 Hilbert 空间, 给定 $B \in \mathcal{L}(Y;Z)$. 证明 $(AB)^* = B^*A^*$.

4.7-2 利用定理 4.7-2 证明: 正规方程组 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}^{\mathrm{T}}\boldsymbol{c}$ (4.4 节) 的解是唯一的, 当且仅当矩阵 $\boldsymbol{A}$ 的秩为 n (由此当然可导出 $n \leqslant m$), 这又等价于当且仅当对称阵 $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$ 为正的.

很自然地, 在复的情况下类似的结论成立.

4.7-3 (Hilbert 空间中的 Lax-Milgram 引理) 设 $(X, (\cdot,\cdot))$ 为实或复的 Hilbert 空间, 映射 $a: X \times X \to \mathbb{K}$ 当 $\mathbb{K} = \mathbb{R}$ 时为双线性型, 当 $\mathbb{K} = \mathbb{C}$ 时对第一个变元是线性的, 对第二个变元是半线性的, 而且存在常数 $M > 0$ 和 $\alpha > 0$, 使得

$$|a(x,y)| \leqslant M\|x\|\|y\|,$$
$$|a(x,x)| \geqslant \alpha\|x\|^2$$

对所有的 $x, y \in X$ 成立, 其中 $\|\cdot\|$ 是由内积 $(\cdot,\cdot)$ 导出的范数.

(1) 证明存在映射 $A \in \mathcal{L}(X)$, 使得对所有的 $x, y \in X$,

$$a(x,y) = (Ax, y).$$

证明这样定义的映射 $A: X \to X$ 是单射, 而且由 $A(X)$ 到 X 上的逆算子是连续的.

---

10) S. Bergman; M. Schiffer [1948]: Kernel functions in the theory of partial differential equations of elliptic type. Duke Mathematical Journal 15, 535–566.

N. Aronszajn; K. T. Smith [1957]: Characterization of positive reproducing kernels. Applications of Green's functions, American Journal of Mathematics, 79, 611–622.

(2) 证明 $A(X)$ 是 X 的闭子空间.

(3) 证明 $A(X) = X$. 由此证明对任意给定的 $b \in X$, 存在唯一的元素 $x \in X$, 使得对一切 $y \in X$,

$$a(x, y) = (b, y)$$

如上定义的映射 $b \in X \to x \in X$ 是连续的.

(3) 的结果构成了 Lax-Milgram 引理 (关于 Lax-Milgram 引理的其他证明将在今后给出, 见定理 6.2-1).

4.7-4 (1) 设 X 和 Y 为两个复 Hilbert 空间 (实的情况是类似的), $A \in \mathcal{L}(X; Y)$, 且 $\operatorname{Im} A$ 是 Y 的闭子空间. 证明存在唯一的 $A^\dagger \in \mathcal{L}(Y; X)$, 称之为 A 的 Moore-Penrose 逆[11], 满足下列四个性质

$$AA^\dagger A = A, \quad A^\dagger A A^\dagger = A^\dagger,$$
$$(AA^\dagger)^* = AA^\dagger, \quad (A^\dagger A)^* = A^\dagger A.$$

(2) 设 $X = \mathbb{C}^n, Y = \mathbb{C}^m$, 此时 A 和 $A^\dagger$ (因 $\operatorname{Im} A$ 是有限维的, 故而必然存在) 可分别恒同于 $m \times n$ 复矩阵 $\boldsymbol{A}$ 和 $n \times m$ 复矩阵 $\boldsymbol{A}^\dagger$. 证明

$$\boldsymbol{A}^\dagger = \lim_{\varepsilon \to 0}((\boldsymbol{A}^*\boldsymbol{A} + \varepsilon I)^{-1}\boldsymbol{A}^*) = \lim_{\varepsilon \to 0}(\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{A}^* + \varepsilon \boldsymbol{I})^{-1}).$$

(3) 证明: 若 $\boldsymbol{A}$ 为 $n \times n$ 复可逆阵, 则 $\boldsymbol{A}^\dagger = \boldsymbol{A}^{-1}$. 这说明了为什么也把 Moore-Penrose 逆 $\boldsymbol{A}^\dagger$ 称之为矩阵 $\boldsymbol{A}$ 的广义逆.

(4) 设 $\boldsymbol{A}$ 为 $m \times n$ 复矩阵. 对任意给定的向量 $\boldsymbol{c} \in \mathbb{C}^m$, 至少存在线性方程组 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{c}$ 的一个最小二乘解 $\boldsymbol{x}$, 即 $\boldsymbol{x}$ 满足 $|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{c}| = \inf_{\boldsymbol{y} \in \mathbb{C}^n} |\boldsymbol{A}\boldsymbol{y} - \boldsymbol{c}|$, 其中 $|\cdot|$ 表示 $\mathbb{C}^m$ 的 Euclid 范数 (定理 4.4-1). 证明存在唯一的向量 $\boldsymbol{x}^\dagger \in \mathbb{C}^n$, 使得

$$|\boldsymbol{x}^\dagger| = \inf\{|\boldsymbol{x}|; \boldsymbol{x} \in \mathbb{C}^n \text{ 且 } |\boldsymbol{A}\boldsymbol{x} - \boldsymbol{c}| = \inf_{\boldsymbol{y} \in \mathbb{C}^n} |\boldsymbol{A}\boldsymbol{y} - \boldsymbol{c}|\},$$

而且这样定义的映射 $\boldsymbol{c} \in \mathbb{C}^m \to \boldsymbol{x}^\dagger \in \mathbb{C}^n$ 恰可表示为 $\boldsymbol{x}^\dagger = \boldsymbol{A}^\dagger \boldsymbol{c}$.

由此可给出矩阵 $\boldsymbol{A}^\dagger$ 的另一个定义.

(5) 设 $\boldsymbol{A}(\varepsilon) = \begin{pmatrix} 1 & 0 \\ 0 & \varepsilon \end{pmatrix}$. 证明 $\lim_{\varepsilon \to 0} \boldsymbol{A}(\varepsilon)^\dagger$ 并不存在[12], 由此证明了矩阵 $\boldsymbol{A}$ 的 Moore-Penrose 逆未必是 $\boldsymbol{A}$ 的元素的连续函数 ($\lim_{\varepsilon \to 0} \boldsymbol{A}(\varepsilon)$ 显然是存在的).

---

[11] E. H. Moore [1920]: On the reciprocal of the general algebraic matrix. Bulletin of the American Mathematical Society 26, 394–395.

R. Penrose [1955]: A generalized inverse for matrices. Proceedings of the Cambridge Philosophical Society 51, 406–413.

这两位作者对这个算子各自提出了两个不同的定义 (在有限的情况下), 其等价性由下面的作者建立:

R. Rado [1956]: Note on generalized inverses of matrices. Proceedings of the Cambridge Philosophical Society 52, 600–601.

各种推广 (在无限维的情况下) 可见:

A. Ben-Israel; T. N. E. Greville [2003]: Generalized Inverses: Theory and Applications. Second Edition, Springer.

[12] 这个例子来自:

G. W. Stewart [1969]: On the continuity of the generalized inverse. Journal of the Society for Industrial and Mathematics 17, 33–45.

4.8 内积空间的极大规范正交系

在 Hilbert 空间中的极大规范正交系起着重要的作用, 正如我们将在下一节就能看到, 这个空间中的任何元素可以展开为这族元素的 Fourier 级数.

如前所述, 设 X 为实 ($\mathbb{K} = \mathbb{R}$) 或复 ($\mathbb{K} = \mathbb{C}$) 的向量空间, 任意给定 $e_i \in X$ 的向量族 $(e_i)_{i\in I}$. 用 $\mathrm{Span}(e_i)_{i\in I}$ 表示由这族向量的所有的有限线性组合构成的 X 的子空间, 即 X 中形如 $\sum_{j\in J}\alpha_j e_j$ 的向量组成的子空间, 其中 J 为 I 的有限子集, 对 $j \in J, \alpha_j \in \mathbb{K}$ (2.1 节).

设 $(X, (\cdot,\cdot))$ 为实或复的内积空间. 由 $e_i \in X$ 组成的元素系 $(e_i)_{i\in I}$ 称之为规范正交系, 是指对所有的 $i, j \in I$ 均有

$$(e_i, e_j) = \delta_{ij}.$$

规范正交系必定是线性独立的元素族, 这是因为给定 I 的有限子族 J, 当 $\sum_{j\in J}\alpha_j e_j = 0$ 时, 对一切 $i \in J$, 有 $\alpha_i = (\sum_{j\in J}\alpha_j e_j, e_i) = 0$.

下面的定理提供了构造规范正交系的简单方法. 为确定起见, 将在无限维空间中作叙述并证明, 有限维空间的情况是显然的.

定理 4.8-1 (Gram-Schmidt[13] 规范正交化方法) 设 $(X, (\cdot,\cdot))$ 为实或复的无限维内积空间, $(f_n)_{n=0}^{\infty}$ 为可列无限个向量 $f_n \in X$ 组成的线性无关向量族. 令

$$\tilde{e}_0 := f_0, \quad \tilde{e}_k = f_k - P_k f_k, \quad k = 1, 2, \cdots,$$

其中 P_k 为 X 到 $\mathrm{Span}(f_n)_{n=0}^{k-1}$ 上的投影算子. 那么, 对所有的 $k \geqslant 1, \tilde{e}_k \neq 0$, 记向量

$$e_n := \frac{\tilde{e}_n}{\|\tilde{e}_n\|}, \quad n \geqslant 0,$$

则向量族 $(e_n)_{n=0}^{\infty}$ 为规范正交系, 它满足对一切 $k \geqslant 0$,

$$\mathrm{Span}(e_n)_{n=0}^{k} = \mathrm{Span}(f_n)_{n=0}^{k},$$

而且

$$\mathrm{Span}(e_n)_{n=0}^{\infty} = \mathrm{Span}(f_n)_{n=0}^{\infty}.$$

---

[13] 这个命名是由于:

J. P. Gram [1883]: Über die Entwicklung reeller Funktionen in in Reihen mittelst der Methode der Kleinsten Quadrate. Journal für die Reine und Angewandte Mathematik 94, 41–73.

E. Schmidt [1907]: Zur Theorie der linearen und nichtlinearen Integralgleichungen. 1. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener, Mathematische Annalen 63, 433–476.

实际上, 规范正交化过程更早已见于:

P. S. Laplace [1820]: Théorie Analytique des Probabilités. Troisième Edition, Premier Supplément: Sur l'Application du Calcul des Probabilités à la Philosophie Naturelle, Courcier, Paris.

证明 因 $\tilde{e}_0 := f_0$, 所以 $\mathrm{Span}(\tilde{e}_0) = \mathrm{Span}(f_0)$. 设对某整数 $k \geqslant 1$, 已取得非零向量 $\tilde{e}_0, \cdots, \tilde{e}_{k-1}$, 满足对所有的 $0 \leqslant m, n \leqslant k-1, m \neq n$,

$$(\tilde{e}_m, \tilde{e}_n) = 0,$$

而且

$$\mathrm{Span}(\tilde{e}_n)_{n=0}^{k-1} = \mathrm{Span}(f_n)_{n=0}^{k-1}.$$

用 P_k 记由 X 到 $X_k := \mathrm{Span}(f_n)_{n=0}^{k-1}$ 上的投影算子 (作为有限维子空间, X_k 在 X 中闭, 见定理 2.7-1(c)). 因为向量 $f_n, 0 \leqslant n \leqslant k$ 线性无关, 故而向量 $\tilde{e}_k := f_k - P_k f_k$ 非零, $\tilde{e}_k$ 正交于子空间 X_k (定理 4.3-1(d)), 从而也正交于所有的向量 $\tilde{e}_n, 0 \leqslant n \leqslant k-1$. 显然, 对一切 $k \geqslant 0, \mathrm{Span}(\tilde{e}_n)_{n=0}^{k} = \mathrm{Span}(f_n)_{n=0}^{k}$, 因此 $\mathrm{Span}(\tilde{e}_n)_{n=0}^{\infty} = \mathrm{Span}(f_n)_{n=0}^{\infty}$.

由此, 定义为 $e_n := \frac{\tilde{e}_n}{\|\tilde{e}_n\|}, n \geqslant 0$, 的向量族 $(e_n)_{n=0}^{\infty}$ 具有所要求的全部性质. □

注 习题 4.8-1 中将利用向量 f_n 给出向量 $\tilde{e}_n$ 的显式表示.

现在我们来给出几个基本的规范正交系的例子. 先考察 (实) 空间 $\mathcal{C}[-1,1]$, 其上赋以空间 $L^2(-1,1)$ 的内积, 即 $(f,g) = \int_{-1}^{1} f(x)g(x)dx$. 对每个整数 $n \geqslant 0$, 函数 $f_n \in \mathcal{C}[-1,1]$ 定义为

$$f_n(x) := x^n, \quad -1 \leqslant x \leqslant 1.$$

按定理 4.8-1, 由向量族 $(f_n)_{n\geqslant 0}$ (它显然是线性无关族) 构造的规范正交系 $(e_n)_{n=0}^{\infty}$ 构成了 Legendre 多项式[14], 其定义为

$$e_n(x) := \frac{\sqrt{n+\frac{1}{2}}}{2^n n!} \frac{d^n}{dx^n}[(x^2-1)^n], \quad -1 \leqslant x \leqslant 1$$

(习题 4.8-2). 注意, 在复空间 $\mathcal{C}([-1,1];\mathbb{C})$ 上赋以空间 $L^2((-1,1);\mathbb{C})$ 的内积, 即 $(f,g) = \int_{-1}^{1} f(x)\overline{g(x)}dx$, 则同样的 Legendre 多项式也组成规范正交系. 自然, Legendre 多项式也是更大的 Hilbert 空间 $L^2(-1,1)$ 或 $L^2((-1,1);\mathbb{C})$ 中的规范正交系.

更一般地, 可以构造紧区间 $[a,b]$ 上的实多项式, 它们关于形为 $(f,g) = \int_a^b f(x)g(x)\omega(x)dx$ 正交, 其中 w 是给定的权函数. 这样的多项式具有一些值得注意的性质[15], 见习题 4.8-3.

下面考察 (实) 空间 $\mathcal{C}_{\mathrm{per}}[0,2\pi]$, 其上赋以空间 $L^2(0,2\pi)$ 的内积, 即 $(f,g) =$

[14] 这个命名是为了纪念 Adrien-Marie Legendre (1752—1833).

[15] 关于多项式的规范正交系的清晰易读的介绍可见 Wong [2010] 和 Beals & Wong [2010]. 这方面最经典的是 Szegö [1975].

$\int_0^{2\pi} f(\theta)g(\theta)d\theta$. 利用初等的三角运算可以验证定义为

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{1}{\sqrt{\pi}}\cos n\theta, \quad \frac{1}{\sqrt{\pi}}\sin n\theta, \quad n \geqslant 1, \quad 0 \leqslant \theta \leqslant 2\pi$$

的函数族组成了空间 $\mathcal{C}_{\text{per}}[0,2\pi]$ 中的规范正交系, 从而也是 Hilbert 空间 $L^2(0,2\pi)$ 中的规范正交系. 类似地, 考察空间 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$, 其上赋以空间 $L^2((0,2\pi),\mathbb{C})$ 的内积, 即 $(f,g) = \int_0^{2\pi} f(\theta)\overline{g(\theta)}d\theta$. 则定义为

$$\frac{1}{\sqrt{2\pi}}e^{in\theta}, \quad n \in \mathbb{Z}, 0 \leqslant \theta \leqslant 2\pi,$$

的函数族组成了空间 $\mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 的规范正交系, 从而也是 Hilbert 空间 $L^2((0,2\pi);\mathbb{C})$ 的规范正交系. 为证明这一点, 只要注意到当 $m \neq n$ 时

$$\int_0^{2\pi} e^{i(m-n)\theta}d\theta = \frac{e^{i(m-n)2\pi}-1}{i(m-n)} = 0,$$

而当 $m = n$ 时

$$\int_0^{2\pi} e^{i(m-n)\theta}d\theta = 2\pi.$$

再考察 (实) Hilbert 空间 $L^2(0,\infty)$, 其上赋以内积 $(f,g) = \int_0^{\infty} f(x)g(x)dx$. 对于 $n \geqslant 0$, 定义为

$$L_n(x) := \frac{1}{n!}e^{\frac{x}{2}}\frac{d^n}{dx^n}[x^n e^{-x}], \quad x \in (0,\infty)$$

的 Laguerre 函数[16] L_n 组成 $L^2(0,\infty)$ 的规范正交系 (习题 4.8-4).

最后考察 (实) Hilbert 空间 $L^2(\mathbb{R})$, 其上赋以内积 $(f,g) = \int_{-\infty}^{\infty} f(x)g(x)dx$. 则对 $n \geqslant 0$, 定义为

$$H_n(x) := \frac{(-1)^n}{2^n n!\sqrt{2\pi}}e^{\frac{x^2}{2}}\frac{d^n}{dx^n}[e^{-x^2}], \quad x \in \mathbb{R},$$

的 Hermite 函数[17] H_n 组成 $L^2(\mathbb{R})$ 的规范正交系 (见习题 4.8-5).

设 $(e_i)_{i\in I}$ 是实或复的内积空间 $(X,(\cdot,\cdot))$ 的规范正交系, 如果满足对所有的 $i \in I$, $(x,e_i) = 0$ 的唯一向量为 $x = 0$, 则称 $(e_i)_{i\in I}$ 为极大的. 经常使用下述的简单的充分条件以讨论极大性.

定理 4.8-2 设 $(e_i)_{i\in I}$ 是内积空间 X 的规范正交系, 如果

$$\overline{\text{Span}(e_i)_{i\in I}} = X,$$

[16] 这个命名是为了纪念 Edmond Laguerre (1834—1886).

[17] 这个命名是为了纪念 Charles Hermite (1822—1901).

则它是极大的[18).]

证明 因为对所有的 $i \in I$ 均有 $(x, e_i) = 0$ 当且仅当 $x \in (\mathrm{Span}(e_i)_{i\in I})^\perp$, 又因为一般地有 $(\overline{W})^\perp = W^\perp$ (定理 4.5-1), 所以规范正交系是极大的当且仅当

$$(\overline{\mathrm{Span}(e_i)_{i\in I}})^\perp = \{0\}.$$

如果 $\overline{\mathrm{Span}(e_i)_{i\in I}} = X$, 则有 $(\overline{\mathrm{Span}(e_i)_{i\in I}})^\perp = \{0\}$, 从而此时 $(e_i)_{i\in I}$ 是极大的. □

值得注意的是正如下面的定理所证明的任何内积空间 (不论完备与否) 均有极大规范正交系. 然而极大规范正交系存在性的证明, 当空间是可分时可采用简单的递推方法 (定理 4.8-3(a)), 否则, 在一般的情况下需要借助于 Zorn 引理或等价地用选择公理 (1.3 节) 做证明 (定理 4.8-4). 为确定起见, 在定理 4.8-3 中我们只讨论无限维的情况 (在有限维情况下, 结论显然同样成立). 我们将在这些空间中给出规范正交系的一个有趣的性质 (见 (b)). 注意, 在 Hilbert 空间中, 定理 4.8-3(a) 的逆命题也成立, 见习题 4.8-6.

注 一般遇到的 Hilbert 空间实际上是可分的 (如 $\ell^2, L^2(\Omega)$ 或 Sobolev 空间 $H^m(\Omega), m \geqslant 1$, 见第 6 章). 但是, 容易构造出不可分的 Hilbert 空间的例子, 见习题 4.8-7.

定理 4.8-3 (可分内积空间中的极大规范正交系) 设 $(X, (\cdot,\cdot))$ 是可分的无限维内积空间.

(a) 存在可数无限个向量 $e_n \in X$ 组成的极大规范正交系 $(e_n)_{n=0}^\infty$, 即满足对所有的 $m, n \geqslant 0$,

$$(e_m, e_n) = \delta_{mn},$$

且对 $x \in X$, 若对所有的 $n \geqslant 0$ 均有 $(x, e_n) = 0$, 必有 $x = 0$.

(b) 任何规范正交系 (不论极大与否) 或为有限的, 或为可数无限的.

证明 因为 X 是可分的, 故而存在线性独立的向量 $f_n \in X, n \geqslant 0$, 使得

$$\overline{\mathrm{Span}(f_n)_{n=0}^\infty} = X$$

(定理 2.2-7). 于是, 利用定理 4.8-1 的 Gram-Schmidt 规范正交化方法, 可由线性无关族 $(f_n)_{n=0}^\infty$ 构造规范正交系 $(e_n)_{n=0}^\infty$, 满足 $\mathrm{Span}(e_n)_{n=0}^\infty = \mathrm{Span}(f_n)_{n=0}^\infty$. 因此

$$\overline{\mathrm{Span}(e_n)_{n=0}^\infty} = \overline{\mathrm{Span}(f_n)_{n=0}^\infty} = X.$$

[18)] 但是, 除非 X 是完备的 (此时, 逆定理易由直和定理导出), 逆命题未必成立: 存在 (必为非完备的) 内积空间 X, 使之没有规范正交系 $(e_i)_{i\in I}$ 满足 $\overline{\mathrm{Span}(e_i)_{i\in I}} = X$. 见:

J. Dixmier [1953]: Sur les bases orthonormales dans les espaces préhibertiens. Acta Scientiarum Mathematicarum Szeged 15, 29–30.

因此, 由定理 4.8-2, $(e_n)_{n=0}^{\infty}$ 是极大的. 这就证明了 (a).

为证明 (b), 首先注意到任何规范正交系 $(e_i)_{i\in I}$ 的元素必定满足当 $i \neq j$ 时

$$\|e_i - e_j\| = \sqrt{\|e_i\|^2 + \|e_j\|^2} = \sqrt{2}.$$

再设向量 $g_k \in X, k \geqslant 0$ 满足 $\overline{\cup_{k=0}^{\infty}\{g_k\}} = X$. 于是, 对每个 $i \in I$, 存在整数 $k(i) \geqslant 0$, 使得 $\|e_i - g_{k(i)}\| \leqslant \frac{\sqrt{2}}{3}$. 这样, 当 $i \neq j$ 时

$$\sqrt{2} = \|e_i - e_j\| \leqslant \|e_i - g_{k(i)}\| + \|g_{k(i)} - g_{k(j)}\| + \|e_j - g_{k(j)}\|,$$

所以 $\|g_{k(i)} - g_{k(j)}\| \geqslant \frac{\sqrt{2}}{3}$, 因为当 $i \neq j$ 时, $g_{k(i)} \neq g_{k(j)}$, 从而 $k(i) \neq k(j)$.

由此, 上面建立的映射 $i \in I \to k(i) \in \mathbb{N}$ 是单射, 从而 I 或为有限的, 或为可数无限的 (1.5 节). □

定理 4.8-4 (任何内积空间中极大规范正交系的存在性) 设 $(X, (\cdot,\cdot))$ 为内积空间, 则存在向量 $e_i \in X$ 组成的向量族 $(e_i)_{i\in I}$, 满足对任何 $i, j \in I$,

$$(e_i, e_j) = \delta_{ij},$$

且对 $x \in X$, 若对所有的 $i \in I$ 均有当 $(x, e_i) = 0$ 时, 必有 $x = 0$.

证明 设 $\dim X \geqslant 2$. 取 $e_1, e_2 \in X$, 使得 $\|e_1\| = \|e_2\| = 1$, 且 $(e_1, e_2) = 0$ (例如, 由两个线性无关的向量 $f_1, f_2 \in X$, 用定理 4.8-3 的方法构造出 e_1 和 e_2).

用 $\mathcal{F}$ 表示由 X 的向量的所有规范正交系组成的 $\mathcal{P}(X)$ 的子集, 它包含 $(e_i)_{i=1,2}$ (从而 $\mathcal{F} \neq \varnothing$), 利用集合的包含关系在其中引入半序 (因为规范正交系 $(e_i)_{i\in I}$ 中的元素互不相同, 所以这里可以恒同于集合 $\cup_{i\in I}\{e_i\}$).

对任意给定的 $\mathcal{F}$ 的全序子集 $\mathcal{E}$, 集合 $G = \cup_{E\in\mathcal{E}} E$ 属于 $\mathcal{F}$. 这是因为如果 $e, \tilde{e} \in G$, 则有某个 $E \in \mathcal{E}$ 使 $e \in E$, 且有某个 $\tilde{E} \in \mathcal{E}$, 使 $\tilde{e} \in \tilde{E}$, 因为 $\mathcal{E}$ 是全序的, 故而或有 $\tilde{E} \subset E$ 或有 $E \subset \tilde{E}$, 从而当 $e \neq \tilde{e}$ 时, $(e, \tilde{e}) = 0$, 当 $e = \tilde{e}$ 时 $(e, \tilde{e}) = 1$ (因 E 和 $\tilde{E}$ 均为规范正交系). 因为对任何 $E \in \mathcal{E}$, 均有 $E \subset G$, 故而 G 是 $\mathcal{E}$ 的一个上界.

由选择公理 (定理 1.3-1), 集合 $\mathcal{F}$ 有极大元 $M = (e_i)_{i\in I}$. 首先, 因 $(e_i)_{i\in I} \in \mathcal{F}$, 故对一切 $i, j \in I, (e_i, e_j) = \delta_{ij}$. 其次, 若 $x \in X$, 使得对一切 $i \in I, (x, e_i) = 0$, 则必有 $x = 0$. 否则, 显然可有集合 $M \cup \left\{\frac{x}{\|x\|}\right\}$ 属于 $\mathcal{F}$, 而 $M \subsetneq M \cup \left\{\frac{x}{\|x\|}\right\}$, 这又与 M 的极大性矛盾. □

要指出的是, 本节前面给出的所有规范正交系都是极大的. 现在我们对前面三个例子给出证明, 而把后两个例子的证明留作练习 (习题 4.8-4 和 4.8-5).

定理 4.8-5 (极大规范正交系的例子) (a) Legendre 多项式, 即

$$e_n(x) = \frac{\sqrt{n + \frac{1}{2}}}{2^n n!} \frac{d^n}{dx^n}[(x^2 - 1)^n], \quad n \geqslant 0, \quad -1 \leqslant x \leqslant 1,$$

组成 Hilbert 空间 $L^2(-1,1)$ 和 $L^2((-1,1);\mathbb{C})$ 的极大规范正交系.

(b) 下列函数:

$$\frac{1}{2\pi},\quad \frac{1}{\pi}\cos m\theta(m\geqslant 1),\quad \frac{1}{\sqrt{\pi}}\sin n\theta(n\geqslant 1),\quad 0\leqslant\theta\leqslant 2\pi,$$

组成 Hilbert 空间 $L^2(0,2\pi)$ 和 $L^2((0,2\pi);\mathbb{C})$ 的极大规范正交系.

(c) 下列函数:

$$\frac{1}{\sqrt{2\pi}}e^{in\theta},\quad n\in\mathbb{Z},\quad 0\leqslant\theta\leqslant 2\pi,$$

组成 Hilbert 空间 $L^2((0,2\pi);\mathbb{C})$ 的极大规范正交系.

证明 (i) 如前所述, 上面这些函数族分别是相应的 Hilbert 空间中的规范正交系. 下面只要证明它们都是极大的.

(ii) 给定函数 $f\in L^2(-1,1)$ 和 $\varepsilon>0$. 因为空间 $\mathcal{C}[-1,1]$ 在空间 $L^2(-1,1)$ 中稠密 (定理 2.5-3), 所以存在函数 $\tilde{f}\in\mathcal{C}[-1,1]$, 使得

$$\|f-\tilde{f}\|_{L^2(-1,1)}\leqslant\frac{\varepsilon}{2},$$

由 Weierstrass 逼近定理 (定理 2.13-3), 存在多项式 p 使得

$$\|\tilde{f}-p\|_{L^2(-1,1)}\leqslant\sqrt{2}\sup_{-1\leqslant x\leqslant 1}|\tilde{f}(x)-p(x)|\leqslant\frac{\varepsilon}{2}.$$

结合这两个不等式即得

$$\overline{\mathrm{Span}(f_n)_{n=0}^{\infty}}=L^2(-1,1),$$

其中 $f_n(x)=x^n, -1\leqslant x\leqslant 1$. 由构造, Legendre 多项式满足

$$\mathrm{Span}(e_n)_{n=0}^{\infty}=\mathrm{Span}(f_n)_{n=0}^{\infty}$$

(定理 4.8-1). 由定理 4.8-2 可知它们构成了空间 $L^2(-1,1)$ 的极大规范正交系.

对空间 $L^2((-1,1);\mathbb{C})$ 中函数的实部与虚部作同样的讨论, 可知 Legendre 多项式组成了空间 $L^2((-1,1);\mathbb{C})$ 的极大规范正交系. 这就证得了 (a).

(iii) 再设给定函数 $g\in L^2(0,2\pi)$ 和 $\varepsilon>0$. 因为空间 $\mathcal{D}(0,2\pi)$ 在 $L^2(0,2\pi)$ 中稠密 (定理 2.6-2), 故而存在函数 $\tilde{g}\in\mathcal{D}(0,2\pi)$ 使得

$$\|g-\tilde{g}\|_{L^2(0,2\pi)}\leqslant\frac{\varepsilon}{2}.$$

因为 $\tilde{g}\in\mathcal{C}_{\mathrm{per}}[0,2\pi]$, 由 Weierstrass 三角多项式的逼近定理 (定理 2.14-3), 存在三角多项式 q, 使得

$$\|\tilde{g}-q\|_{L^2(0,2\pi)}\leqslant\sqrt{2\pi}\sup_{0\leqslant\theta\leqslant 2\pi}|\tilde{g}(\theta)-q(\theta)|\leqslant\frac{\varepsilon}{2}.$$

这就证得了 (b).

(iv) (c) 的证明类似于 (b), 只是其中应用 Weierstrass 三角多项式逼近定理之处被其复数形式 (定理 2.15-4) 所取代后断言对任意给定的函数 $\hat{g} \in \mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$, 存在复三角多项式 q, 即

$$q(\theta) = \sum_{k=-n}^{n} c_k e^{ik\theta}, \quad 0 \leqslant \theta \leqslant 2\pi,$$

其中 $c_k \in \mathbb{C}, -n \leqslant k \leqslant n, n \geqslant 0$, 关于 $[0,2\pi]$ 的上确界范数, q 可任意地逼近于 $\hat{g}$. □

注 很自然地, Legendre 多项式对于 $L^2(-1,1)$ 中任何包含它们的子空间而言, 也构成极大规范正交系, 这样的子空间如 $\mathcal{C}[-1,1]$ (赋以 $L^2(-1,1)$ 的内积).

我们在后面将会看到, 可分 Hilbert 空间上紧自伴算子的谱定理 (4.11 节) 也提供了另一个在这类空间中构造极大规范正交系的有效途径 (定理 4.11-3). 一个基本而特殊的例子, 求解二阶椭圆型算子的特征值问题将在后面介绍 (定理 6.10-2).

习题

4.8-1 证明由 Gram-Schmidt 规范正交化方法 (定理 4.8-1) 构造的向量 $\tilde{e}_n, n \geqslant 1$, 可以递推地表示为 $\tilde{e}_0 = f_0$, 当 $n \geqslant 1$ 时, $\tilde{e}_n := f_n - \sum_{k=0}^{n-1} \frac{(f,\tilde{e}_k)}{\|\tilde{e}_k\|^2}\tilde{e}_k$.

4.8-2 (1) 对每个整数 $n \geqslant 0$, 函数 $f_n : [-1,1] \to \mathbb{R}$ 定义为 $f_n(x) := x^n, -1 \leqslant x \leqslant 1$. 证明由函数族 $(f_n)_{n=0}^{\infty}$ 按定理 4.8-1 构造的规范正交系 $(e_n)_{n=0}^{\infty}$ 即由 n 次 Legendre 多项式组成, 即对一切 $n \geqslant 0$,

$$e_n(x) := \frac{\sqrt{n+\frac{1}{2}}}{2^n n!}\frac{d^n}{dx^n}[(x^2-1)^m], \quad -1 \leqslant x \leqslant 1.$$

(2) 直接验证对所有的 $m, n \geqslant 0$, $\displaystyle\int_{-1}^{1} e_m(x)e_n(x)dx = \delta_{mn}$.

(3) 直接验证对所有的 $n \geqslant 0$,

$$e_n(x) = \sqrt{n+\frac{1}{2}}\sum_{0\leqslant k\leqslant\frac{n}{2}}(-1)^k\frac{(2(n-k))!}{2^k(n-k)!k!(n-2k)!}x^{n-2k}, \quad -1 \leqslant x \leqslant 1,$$

特别地, 这说明 e_n 是 n 次多项式 (这又可由定理 4.8-1 得到).

(4) 证明: 反之, 对每个 $n \geqslant 0$,

$$x^n = n!\sum_{0\leqslant k\leqslant\frac{n}{2}}\left(\frac{1}{(2k)!\sqrt{n-2k+\frac{1}{2}}}\right)\left(\frac{2n-4k+1}{(2k+1)(2k+3)\cdot(2n-2k+1)}\right)e_{n-2k}(x), \quad -1 \leqslant x \leqslant 1.$$

(5) 设二阶微分算子 $\mathcal{L}$ 定义为

$$\mathcal{L}u(x) := -\frac{d}{dx}\left[(1-x^2)\frac{du}{dx}\right], \quad -1 \leqslant x \leqslant 1.$$

证明: 对每个 $n \geqslant 0$, Legendre 多项式 e_n 是算子 $\mathcal{L}$ 的特征函数, 即 e_n 满足

$$\mathcal{L}e_n(x) = \lambda_n e_n(x), \quad -1 \leqslant x \leqslant 1,$$

其中 $\lambda_n = n(n+1)$.

4.8-3 (关于权函数的正交多项式) 设 ω 为区间 $[0,1]$ 上的一个权函数, 即函数 $\omega \in L^1(0,1)$, 且在 $[0,1]$ 上几乎处处 $\omega > 0$. 则

$$(f,g) := \int_0^1 f(x)g(x)\omega(x)dx$$

显然定义了空间 $\mathcal{C}[0,1]$ 上的一个内积.

在本题中, "正交" 或 "规范正交" 均关于这个内积. 对每个 $n \geqslant 0$, 令 $f_n(x) := x^n, 0 \leqslant x \leqslant 1$, 设 $(e_n)_{n=0}^{\infty}$ 是由向量族 $(f_n)_{n=0}^{\infty}$ 按定理 4.8-1 的方法构造而得的规范正交系. 因为对所有的 $k \geqslant 0$, $\mathrm{Span}(e_n)_{n=0}^k = \mathrm{Span}(f_n)_{n=0}^k$, 对 $n \geqslant 0$, 每个多项式 e_n 形为 $e_n(x) = c_n x^n + c_{n-1}x^{n-1} + \cdots + c_0, 0 \leqslant x \leqslant 1$, 其中 $c_n \neq 0$. 对 $n \geqslant 0$, 多项式 $p_n \in \mathcal{P}_n[0,1]$ 定义为

$$p_n(x) := \frac{1}{c_n}e_n(x) = x^n + \frac{c_{n-1}}{c_n}x^{n-1} + \cdots + \frac{c_0}{c_n}, \quad 0 \leqslant x \leqslant 1,$$

称之为关于权函数 ω 的 monic 正交多项式 ("monic" 意即首项系数为 1). 注意, 这些多项式满足当 $m \neq n$ 时 $(p_m, p_n) = 0$, 但不再满足对任何 $n \geqslant 0, (p_n, p_n) = 1$. 本题的目的在于建立这些多项式的两个基本性质.

(1) 证明多项式 p_n 满足下述的三项递推公式, 即对 $n \geqslant 2$,

$$p_n(x) = (x + b_n)p_{n-1}(x) + c_n p_{n-2}(x), \quad 0 \leqslant x \leqslant 1,$$

其中常数 $b_n, c_n \in \mathbb{R}$ 是多项式 p_n, p_{n-1}, p_{n-2} 系数的函数.

(2) 证明对每个 $n \geqslant 1$, p_n 作为 $\mathbb{R}$ 上的多项式, 其根均为实的单根, 而且全在开区间 $]0,1[$ 中.

4.8-4 (1) 对每个 $n \geqslant 0$, 设函数 $f_n : [0,\infty[\to \mathbb{R}$ 定义为 $f(x) := e^{-\frac{x}{2}}x^n, x \in [0,\infty[$. 证明函数 f_n 属于空间 $L^2(0,\infty)$, 而且由函数族 $(f_n)_{n=0}^{\infty}$ 按定理 4.8-1 的方法构得的规范正交系 (它显然是线性无关的) 即由 Laguerre 函数 $L_n, n \geqslant 0$, 组成, 其定义为

$$L_n(x) := \frac{1}{n!}e^{\frac{x}{2}}\frac{d^n}{dx^n}[x^n e^{-x}], \quad x \in [0,\infty[.$$

(2) 证明在 Hilbert 空间 $L^2(0,\infty)$ 中, 规范正交系 $(L_n)_{n=0}^{\infty}$ 是极大的[19].

4.8-5 (1) 对每个 $n \geqslant 0$, 函数 $f_n : \mathbb{R} \to \mathbb{R}$ 定义为 $f_n(x) := e^{-\frac{x^2}{2}}x^n, x \in \mathbb{R}$. 证明函数 f_n 属于空间 $L^2(\mathbb{R})$, 且由函数族 $(f_n)_{n=0}^{\infty}$ (它们显然是线性无关的) 按定理 4.8-1 构造的规范正交系即由 Hermite 函数 H_n 组成, $n \geqslant 0$, 其定义为

$$H_n(x) := \frac{(-1)^n}{2^n n! \sqrt{2\pi}}e^{\frac{x^2}{2}}\frac{d^n}{dx^n}[e^{-x^2}], \quad x \in \mathbb{R}.$$

(2) 证明在 Hilbert 空间 $L^2(\mathbb{R})$ 中, 规范正交系 $\{H_n\}$ 是极大的[20].

---

[19] 证明可见 Goffman; Pedrick [1965, 4.10 节].

[20] 证明可见 Akhiezer; Glazman [1961, 11 节].

4.8-6 设 X 为实或复的 Hilbert 空间, 且有有限或可列的极大的规范正交基. 证明 X 是可分的.

4.8-7 本题提供了一个不可分的 Hilbert 空间的例子, 且它有一个不可数的规范正交系.

(1) 设复向量空间 $\mathcal{C}(\mathbb{R};\mathbb{C})$ 的子空间 Y 定义为

$$Y := \mathrm{Span}(e_\lambda)_{\lambda\in\mathbb{R}},$$

其中

$$e_\lambda(x) := e^{i\lambda x}, \quad x \in \mathbb{R}.$$

证明 $(f,g) := \lim_{T\to\infty} \frac{1}{2T}\displaystyle\int_{-T}^{T} f(x)\overline{g(x)}dx$ 定义了 Y 上的一个内积.

(2) 证明 $(e_\lambda)_{\lambda\in\mathbb{R}}$ 是空间 $(Y,(\cdot,\cdot))$ 上的规范正交系.

(3) 证明 $(e_\lambda)_{\lambda\in\mathbb{R}}$ 是空间 $(Y,(\cdot,\cdot))$ 的极大规范正交系.

(4) 证明 Y 的完备化空间 X 是一个不可分的 Hilbert 空间 (定理 4.1-4).

4.9 Hilbert 空间中的 Hilbert 基和 Fourier 级数

在上一节中, 我们已经看到在任何无限维的内积空间 X 中极大规范正交系 $(e_i)_{i\in I}$ 总是存在的, 如果 X 是可分的, 则它是可数无限的 (定理 4.8-3 和 4.8-4). 现在我们来证明, 如果 X 是完备的, 则这样的向量族具有一个基本性质, 即任何元素 $x \in X$ 可以展开为形为 $\sum_{i\in I}(x,e_i)e_i$ 的级数. 由于这个原因, Hilbert 空间 X 中的极大规范正交系被称之为 X 的 Hilbert 基.

下面的定理是线性泛函分析最基本的结果之一. 这里仅考察可分的情况 (习题 4.8-6), 而把可分情况下的若干补充和不可分的情况留作习题 (习题 4.9-1 和 4.9-2).

定理 4.9-1 (可分 Hilbert 空间的 Fourier 级数) 设 $(X,(\cdot,\cdot))$ 为无限维 Hilbert 空间, $(e_n)_{n=1}^\infty$ 为 X 的一个 Hilbert 基.

(a) 任何元素 $x \in X$ 均可展开为收敛级数

$$x = \sum_{n=1}^{\infty}(x,e_n)e_n,$$

它称之为 x 的 Fourier 级数[21].

(b) 对 $n \geqslant 1$, 数 $(x,e_n) \in \mathbb{K}$ 称之为 x (关于基 $(e_n)_{n=1}^\infty$) 的 Fourier 系数, 它满足 Parseval 公式[22]:

---

[21] 这样命名是为了纪念 Jean Baptiste Joseph Fourier (1768—1830) 和他关于热导理论的重要著作: 1822 年出版的 *Théorie Analytique de la Chaleur*. 在这一杰作中, Fourier 建立了 “经典” Fourier 级数 (即利用正弦和余弦函数展开, 这些级数将在本节后面部分定义) 在某些特殊情况下的收敛性, 并且说明了如何使用 Fourier 级数来解偏微分方程, 如热导方程.

[22] 这样命名是为了纪念 Marc-Antoine Parseval, 他在 1799 年通过直接计算推导出经典 Fourier 级数的系数 a_k 和 b_k 满足 $\frac{1}{2}\int_0^{2\pi}|g(\theta)|^2d\theta = \frac{a_0^2}{2} + \sum_{k=1}^\infty(a_k^2+b_k^2)$, 但是尚无收敛性的证明.

$$\|x\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2.$$

证明 假设给定 $x \in X$.

(i) 首先证明 $\sum_{n=1}^{\infty} |(x, e_n)|^2 < \infty$. 利用假设 $(e_n)_{n=1}^{\infty}$ 是规范正交系, 可得对任何整数 $k \geqslant 1$,

$$\begin{aligned}
0 &\leqslant \left\| x - \sum_{n=1}^{k} (x, e_n) e_n \right\|^2 \\
&= \left(x - \sum_{n=1}^{k} (x, e_n) e_n, x - \sum_{m=1}^{k} (x, e_m) e_m \right) \\
&= \|x\|^2 - \sum_{n=1}^{k} |(x, e_n)|^2 - \sum_{m=1}^{k} |(x, e_m)|^2 + \sum_{m,n=1}^{k} (x, e_n)\overline{(x, e_m)}(e_n, e_m) \\
&= \|x\|^2 - \sum_{n=1}^{k} |(x, e_n)|^2.
\end{aligned}$$

由这个不等式即得对任意的 $k \geqslant 1$,

$$\sum_{n=1}^{k} |(x, e_n)|^2 \leqslant \|x\|^2,$$

由此可知级数 $\sum_{n=1}^{\infty} |(x, e_n)|^2$ 收敛.

(ii) 再证明在空间 X 中, $\sum_{n=1}^{\infty} (x, e_n) e_n$ 是收敛级数 (3.6 节). 因为 X 是完备的, 只要证明由

$$x_k := \sum_{n=1}^{k} (x, e_n) e_n$$

所定义而得的序列 $(x_k)_{k=1}^{\infty}$ 是 Cauchy 序列. 为此, 再次利用族 $(e_n)_{n=1}^{\infty}$ 是规范正交的, 于是对任何整数 $\ell \geqslant 1$ 和 $k \geqslant \ell + 1$, 有

$$\begin{aligned}
\|x_k - x_\ell\|^2 &= \left(\sum_{n=\ell+1}^{k} (x, e_n) e_n, \sum_{m=\ell+1}^{k} (x, e_m) e_m \right) \\
&= \sum_{n=\ell+1}^{k} |(x, e_n)|^2.
\end{aligned}$$

由 (i), 级数 $\sum_{n=1}^{\infty} |(x, e_n)|^2$ 收敛, 所以 $(x_k)_{k=1}^{\infty}$ 为 Cauchy 序列.

(iii) 令 $y := \lim\limits_{k \to \infty} x_k = \sum_{n=1}^{\infty} (x, e_n) e_n$. 余下要证明 $x = y$; 或者因为由假设, 规范正交系 $(e_n)_{n=1}^{\infty}$ 是极大的, 又等价于证明对任何 $n \geqslant 1$, 均有

$$(x - y, e_n) = 0.$$

由 y 的定义和内积的连续性可得

$$(x-y,e_n)=\lim_{k\to\infty}\left(x-\sum_{m=1}^{k}(x,e_m)e_m,e_n\right).$$

而当 $k\geqslant n$ 时

$$\left(x-\sum_{m=1}^{k}(x,e_m)e_m,e_n\right)=0.$$

因此 $x=y$.

(iv) 由 (i) 中出现的关系式和 (ii), (iii) 中建立的关系式 $x=\lim_{k\to\infty}\sum_{n=1}^{k}(x,e_n)e_n$, 即得

$$0=\lim_{k\to\infty}\left\|x-\sum_{n=1}^{k}(x,e_n)e_n\right\|^2=\lim_{k\to\infty}\left(\|x\|^2-\sum_{n=1}^{k}|(x,e_n)|^2\right).$$

这就证得了 Parseval 公式. □

由上述证明中的部分 (i) 可见, 在任何内积空间 X 中 (不论完备与否), 对任何 $x\in X$ 和任何规范正交系 $(e_n)_{n=1}^{\infty}$ (不论极大与否), 不等式

$$\sum_{n=1}^{\infty}|(x,e_n)|^2\leqslant\|x\|^2$$

总是成立的. 这个不等式称之为 Bessel 不等式[23].

注意, 由级数 $\sum_{n=1}^{\infty}|(x,e_n)|^2$ 的收敛性 (Bessel 不等式的推论); 显然可导出对任何内积空间 X 中的任何规范正交系 $(e_n)_{n=1}^{\infty}$ 和任何 $x\in X$,

$$\lim_{n\to\infty}(x,e_n)=0.$$

定理 4.9-1 有许多重要的推论. 例如, 将它应用于空间 $L^2(0,2\pi)$, 可以导出任何 (实) 函数 $g\in L^2(0,2\pi)$ 均可用在定理 4.8-5(b) 中定义的 Hilbert 基展开为 "经典" 的 Fourier 级数 ("经典的" 是相对于以定理 4.9-1 中考察的任意 Hilbert 基作展开的 "一般的" Fourier 级数而言).

定理 4.9-2 (经典 Fourier 级数) 给定任意函数 $g\in L^2(0,2\pi)$, 对任何 $n\geqslant 0, g$ 的 n 次 Fourier 部分和 $S_n g\in\mathcal{C}_{\text{per}}[0,2\pi]$ 定义为

$$(S_n g)(\theta):=\frac{a_0}{2}+\sum_{k=1}^{n}(a_k\cos k\theta+b_k\sin k\theta), 0\leqslant\theta\leqslant 2\pi,$$

[23] 这个命名是为了纪念 Friedrich Wilhelm Bessel (1784—1846), 他于 1828 年对于 "经典" Fourier 级数的系数建立了这个不等式.

其中

$$a_k := \frac{1}{\pi} \int_0^{2\pi} g(\varphi) \cos k\varphi d\varphi, \quad k \geqslant 0,$$
$$b_k := \frac{1}{\pi} \int_0^{2\pi} g(\varphi) \sin k\varphi d\varphi, \quad k \geqslant 1.$$

则当 $n \to \infty$ 时

$$\|S_n g - g\|_{L^2(0,2\pi)} \to 0,$$

相应的经典 Fourier 级数的 Parseval 公式为

$$\|g\|^2_{L^2(0,2\pi)} = \pi \left(\frac{|a_0|^2}{2} + \sum_{k=1}^{\infty} |a_k|^2 + \sum_{k=1}^{\infty} |b_k|^2 \right).$$

由这个定理可知, 存在子列 $(S_{\sigma(n)}g)_{n=1}^{\infty}$ 在区间 $[0,2\pi]$ 上几乎处处收敛于 g (定理 3.4-3). 在 1913 年提出的 Lusin 猜想[24] 是实际上整个序列 $(S_n g)_{n=1}^{\infty}$ 应在 $[0,2\pi]$ 上几乎处处收敛. 这个看似平淡无奇的命题在几十年中成为最具挑战性的问题之一, 直到 1966 年才由 Lennart Carleson 在一篇里程碑式的论文[25] 中证明是正确的.

注 即使函数 g 是 $[0,2\pi]$ 上连续的周期函数, 其 Fourier 级数未必在 $[0,2\pi]$ 上一致收敛: 后面我们将说明这个断言是 Banach-Steinhaus 定理 (定理 5.5-1) 的推论. 不同的是如果用 F_n 表示 Fejér 算子, 则三角多项式 $F_n g$ 一致收敛于 g (定理 2.14-2).

类似地, 任何复值函数 $g \in L^2((0,2\pi);\mathbb{C})$ 可以用定理 4.8-4(c) 中定义的 Hilbert 基展开为 Fourier 级数:

定理 4.9-3 (复数情况下的经典 Fourier 级数) 任意给定函数 $g \in L^2((0,2\pi);\mathbb{C})$, 对任何 $n \geqslant 0$, n 次 Fourier 部分和 $g_n \in \mathcal{C}_{\text{per}}([0,2\pi];\mathbb{C})$ 定义为

$$g_n(\theta) = \sum_{k=-n}^{n} c_k e^{ik\theta}, \quad 0 \leqslant \theta \leqslant 2\pi,$$

其中

$$c_k := \frac{1}{2\pi} \int_0^{2\pi} g(\varphi) e^{-ik\varphi} d\varphi, \quad k \geqslant 0.$$

则当 $n \to \infty$ 时

$$\|g_n - g\|_{L^2((0,2\pi);\mathbb{C})} \to 0,$$

---

[24] N. Lusin [1913]: Sur la convergence des séries trigonométriques de Fourier. Comptes Rendus de l'Académie des Sciences de Paris 156, 1655–1658.

[25] L. Carleson [1966]: On convergence and growth of partial sums of Fourier series. Acta Mathematica 116, 135–157.

由于这项工作和其他的数学成果, Carleson 于 2006 年获得了 Abel 奖.

相应的经典复 Fourier 级数的 Parseval 公式为

$$\|g\|^2_{L^2((0,2\pi);\mathbb{C})} = 2\pi \sum_{k=-\infty}^{\infty} |c_k|^2.$$

注 定理 4.9-2 中的系数 a_k 和 b_k, 以及定理 4.9-3 中定义的系数 c_k, 并非定理 4.9-1 中定义的真正的 Fourier 系数; 与后者相应的应该为 $\sqrt{\frac{\pi}{2}}a_0, \sqrt{\pi}a_k, \sqrt{\pi}b_k, k \geqslant 1$, 和 $\sqrt{2\pi}c_k, k \in \mathbb{Z}$. 这就说明为什么在相应的 Parseval 公式中出现因子 π 和 2π.

注意, 将 Fourier 系数收敛于 0 的性质应用到上面的例子, 则可导出对任意函数 $g \in L^2(0, 2\pi)$,

$$\lim_{n\to\infty} \int_0^{2\pi} g(\varphi)\cos n\varphi d\varphi = 0, \quad \lim_{n\to\infty} \int_0^{2\pi} g(\varphi)\sin n\varphi d\phi = 0,$$

对任何函数 $g \in L^2((0, 2\pi); \mathbb{C})$,

$$\lim_{n\to\infty} \int_0^{2\pi} g(\theta)e^{-in\theta}d\theta = 0, \quad \lim_{n\to-\infty} \int_0^{2\pi} g(\theta)e^{-in\theta}d\theta = 0.$$

这些关系式构成了 Riemann-Lebesgue 引理, 它们为第 5 章中讨论的一个基本概念, 即弱收敛, 提供了例子.

将定理 4.9-1 应用于空间 $L^2((0, 2\pi); \mathbb{C})$ 得到的另一个重要结论是 Riesz-Fischer 定理 (见习题 4.9-4).

自然地, 在空间 $L^2(-1, 1), L^2((-1, 1); \mathbb{C}), L^2(0, \infty), L^2(\mathbb{R})$ 中, 相应地利用 Legendre 多项式, Laguerre 函数和 Hermite 函数可得类似的 Fourier 展开式和 Parseval 公式.

回到一般的情况. 再次利用定理 4.9-1 可知, 在任何可分 Hilbert 空间 X 和 ℓ^2 之间, 存在 Hilbert 空间的同构: 即存在 X 与 ℓ^2 间的线性双射 (在下面的定理中记作 σ), 它保持内积 (从而是等距的), 因此, 两者的 Hilbert 空间结构是相同的.

定理 4.9-4 设 $(X, (\cdot, \cdot))$ 为实或复的无限维可分 Hilbert 空间, 则存在从 X 到实或相应的复空间 ℓ^2 上的线性双射 σ, 使得对任何 $x, y \in X$,

$$(x, y)_X = (\sigma x, \sigma y)_{\ell^2}.$$

由此, 借助于保持内积的线性等距, 任何无限维可分 Hilbert 空间恒同于空间 ℓ^2.

证明 因为 X 是无限维可分 Hilbert 空间, 故而在 X 中存在可数无限的 Hilbert 基 $(e_n)_{n=1}^{\infty}$ (定理 4.8-3). 因此, 任何 $x \in X$ 可以展开为 Fourier 级数 $x = \sum_{n=1}^{\infty}(x, e_n)e_n$ (定理 4.9-1) 于是, 对每个 $x \in X$, 令

$$\sigma(x) := ((x, e_n))_{n=1}^{\infty}.$$

首先, 对每个 $x \in X$. 由 Parseval 公式 (定理 4.9-1), $\|\sigma(x)\|_{\ell^2} = \sum_{n=1}^{\infty} |(x, e_n)|^2 = \|x\|^2 < \infty$, 所以 $\sigma(x) \in \ell^2$.

其次, 这样定义的映射 $\sigma : X \to \ell^2$ 是线性的 (内积关于第一个变量是线性的), 也是等距的 (由 Parseval 公式), 而且保持内积, 这是因为

$$
\begin{aligned}
(x, y)_X &= \lim_{k \to \infty} \left(\sum_{m=1}^{k} (x, e_m) e_m, \sum_{n=1}^{k} (y, e_n) e_n \right) \\
&= \lim_{k \to \infty} \sum_{n=1}^{k} (x, e_n) \overline{(y, e_n)} = (\sigma(x), \sigma(y))_{\ell^2}
\end{aligned}
$$

(利用内积的连续性, 见定理 4.1-1(c)).

最后, 给定任意的元素 $\xi = (\xi_n)_{n=1}^{\infty} \in \ell^2$ 和任意的整数 $k \geqslant 1$, 令 $x_k := \sum_{n=1}^{k} \xi_n e_n$. 则因为对所有的 $k - 1 \geqslant \ell \geqslant 1$, 均有

$$
\|x_k - x_\ell\|^2 = \sum_{n=\ell+1}^{k} |\xi_n|^2.
$$

而由假设 $\sum_{n=1}^{\infty} |\xi_n|^2 < \infty$, 所以序列 $(x_k)_{k=1}^{\infty}$ 为 X 中的 Cauchy 序列. 设 $x := \lim_{k \to \infty} x_k$ (由假设, X 是完备的). 于是, 对每个 $n \geqslant 1$,

$$
(x, e_n) = \lim_{k \to \infty} (x_k, e_n) = \xi_n.
$$

因此, $\sigma(x) = \xi$, 这就证明了 $\sigma : X \to \ell^2$ 是满射. 从而映射 σ 具备要证明的全部性质. □

习题

4.9-1 设 $(X, (\cdot, \cdot))$ 为无限维可分 Hilbert 空间, $(e_n)_{n=1}^{\infty}$ 是 X 中的规范正交系. 证明下列性质等价:

(1) 向量族 $(e_n)_{n=1}^{\infty}$ 是极大的.

(2) 任何元素 $x \in X$ 可展开为 $x = \sum_{n=1}^{\infty} (x, e_n) e_n$.

(3) 对任何 $x \in X$, $\|x\|^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2$.

(4) 对任何 $x, y \in X$, $(x, y) = \sum_{n=1}^{\infty} (x, e_n) \overline{(y, e_n)}$.

(5) $\overline{\mathrm{Span}(e_n)_{n=1}^{\infty}} = X$.

注 定理 4.9-1 已经证明由 (1) 可以导出 (2) 和 (3).

4.9-2 (不可分 Hilbert 空间中的 Fourier 级数) 本题构成了定理 4.9-1 的“不可分形式”.

(1) 设 $(X, (\cdot, \cdot))$ 为内积空间, $(e_i)_{i \in I}$ 是元素 $e_i \in X$ 组成的不可数无限的规范正交系. 证明对任意给定的 $x \in X$, 至多只有可数多个足标 $i \in I$, 使得 $(x, e_i) \neq 0$.

(2) 设 $(X,(\cdot,\cdot))$ 为不可分 Hilbert 空间, $(e_i)_{i\in I}$ 是 X 的 Hilbert 基 (由定理 4.8-4, 这样的 Hilbert 基存在, 由习题 4.8-6, 它必是不可数无限的). 任意给定 $x\in X$, 把 $((x,e_i))_{i\in I}$ 中的非零数排列为序列 $(x,e_n)_{n=0}^{\infty}$ ($(x,e_i)_{i\in I}$ 中只有有限个数非零的情况留给读者). 证明 $x=\sum_{n=0}^{\infty}(x,e_n)e_n$, 这个级数是可交换收敛的, 即对任何双射 $\gamma:\mathbb{N}\to\mathbb{N}$, $x=\sum_{n=0}^{\infty}(x,e_{\gamma(n)})e_{\gamma(n)}$.

(3) 证明 $\|x\|^2=\sum_{n=0}^{\infty}|(x,e_n)|^2$, 级数 $\sum_{n=0}^{\infty}|(x,e_n)|^2$ 是可交换收敛的.

4.9-3 设 X 为 Hilbert 空间, 证明 X 的任何两个 Hilbert 基具有相同的基数 (1.5 节).

4.9-4 (F. Riesz-Fischer 定理[26]) 给定 $c_k\in\mathbb{C}, k\in\mathbb{Z}$, 使得 $\sum_{k=-\infty}^{\infty}|c_k|^2<\infty$. 证明存在函数 $g\in L^2((0,2\pi);\mathbb{C})$. 使得对所有的 $k\in\mathbb{Z}, c_k=\dfrac{1}{2\pi}\displaystyle\int_0^{2\pi}g(\varphi)e^{-ik\varphi}d\varphi$.

4.9-5 设 G 为空间 $L^2(]0,1[\times]0,1[)$ 中的函数.

(1) 任意给定函数 $f\in L^2(0,1)$, 令

$$Af(x):=\int_0^1 G(x,\xi)f(\xi)d\xi,\quad 0\leqslant x\leqslant 1.$$

证明这个关系式定义了一个函数 $Af\in L^2(0,1)$, 由此定义的线性算子 $A:L^2(0,1)\to L^2(0,1)$ 是紧的.

提示: 在空间 $L^2(0,1)$ 中取 Hilbert 基 $(e_n)_{n=1}^{\infty}$, 对每个 $n\geqslant 1$, 定义线性算子 $A_n: L^2(0,1)\to \mathrm{Span}(e_k)_{k=1}^{n}$ 为对任何 $f\in L^2(0,1)$, $A_nf=\sum_{\ell,k=1}^{m}(Ae_\ell,e_k)(f,e_\ell)e_k$. 证明当 $n\to\infty$ 时 $\|A_n-A\|\to 0$, 再利用习题 3.2-4.

(2) 如果对几乎所有的 $(x,\xi)\in[0,1]\times[0,1]$, $G(x,\xi)=G(\xi,x)$, 证明算子 A 对任何 $f,g\in L^2(0,1)$ 满足 $(Af,g)=(f,Ag)$, 其中 $(\cdot,\cdot)$ 为空间 $L^2(0,1)$ 的内积.

4.10 内积空间中的自伴算子的特征值和特征函数

设 $(X,(\cdot,\cdot))$ 为 $\mathbb{K}$ 上的内积空间. 如果线性算子 $A:X\to X$ 和它的伴随算子 A^* (4.7 节) 相等, 即对任何 $x,y\in X$,

$$(Ax,y)=(x,Ay),$$

则称 A 为自伴算子. 自伴算子当 $\mathbb{K}=\mathbb{R}$ 时也称作对称算子, 当 $\mathbb{K}=\mathbb{C}$ 时也称作 Hermite 算子.

注 后面将会看到 (定理 5.7-2), 如果 X 是 Hilbert 空间, 则由 X 到 X 的任何自伴算子都连续. 这个值得注意的, 甚至有些惊人的性质其实是 Banach 闭图像定理的简单推论.

---

[26] 这个定理首先发表于:

F. Riesz [1907]: Sur les systèmes orthogonaux de fonctions. Comptes Rendus de l'Académie des Sciences 144, 615–619.

几乎立刻给出了另一个证明:

E. Fischer [1907]: Sur la convergence en moyenne. Comptes Rendus de l'Acaeémie des Sciences 144, 1022–1024.

例如, 设 $\mathbb{R}^n$ 上赋以 Euclid 内积 (4.2 节) 因为从 $\mathbb{R}^n$ 到 $\mathbb{R}^n$ 中的任何线性算子都可恒同于一个 $n \times n$ 实矩阵 $\boldsymbol{A}$, 显然可得这样的线性算子为对称的充要条件是相应的矩阵 $\boldsymbol{A} = (a_{ij})$ 是对称阵, 即对所有的 $1 \leqslant i, j \leqslant n$, 有 $a_{ij} = a_{ji}$.

类似地, 设 $\mathbb{C}^n$ 上赋以 Hermite 内积 (4.2 节). 因为从 $\mathbb{C}^n$ 到 $\mathbb{C}^n$ 中任何线性算子都可恒同于一个 $n \times n$ 复矩阵 $\boldsymbol{A}$, 同样显然可得这样的线性算子为 Hermite 的充要条件是相应的矩阵 $\boldsymbol{A} = (a_{ij})$ 为 Hermite 阵, 即对所有的 $1 \leqslant i, j \leqslant n$, 有 $a_{ij} = \bar{a}_{ji}$.

设 $A : X \to X$ 为自伴线性算子, 如果对所有的 $x \in X$ 均有 $(Ax, x) \geqslant 0$, 则称 A 为*非负定的*, 如果对所有非零的 $x \in X$ 均有 $(Ax, x) > 0$, 则称 A 为*正定的*. 注意, 如果 A 是正定的, 则 $\operatorname{Ker} A = \{0\}$.

一般的自伴算子的非负定和正定的概念, 推广了对实对称阵或复 Hermite 阵所熟知的相应概念.

从空间 $(\mathcal{C}[0,1], (\cdot,\cdot))$ 到其自身或从空间 $(L^2(0,1), (\cdot,\cdot))$ 到其自身的对称线性算子的例子分别在习题 3.10-4 和 4.9-5 中给出, 在这两种情况下, $(\cdot,\cdot)$ 均表示空间 $L^2(0,1)$ 的内积.

下面的定理列出了自伴算子特征值和特征向量的一系列初等而又常用的性质, 它们推广了实对称阵和复 Hermite 阵的一些熟知的性质.

定理 4.10-1 *设 $(X, (\cdot,\cdot))$ 为内积空间, $A : X \to X$ 为自伴线性算子.*

(a) 对任何 $x \in X$, 数 (Ax, x) 是实的.

(b) 设 λ 为 A 的任意的特征值, 则 λ 必是实数. 且若 A 是非负定的, 则 $\lambda \geqslant 0$; 若 A 是正定的, 则 $\lambda > 0$.

(c) 对应于不同特征值的特征向量相互正交.

(d) 如果 $A \in \mathcal{L}(X)$, A 的算子范数, 即 $\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$, 也可由

$$\|A\| = \sup_{x \neq 0} \frac{(Ax, x)}{\|x\|^2}$$

给出.

证明 如果 $\mathbb{K} = \mathbb{C}$, 则对任何 $x \in X, (Ax, x) = (x, Ax) = \overline{(Ax, x)}$. 因此, $(Ax, x) \in \mathbb{R}$. 这就证得 (a) ($\mathbb{K} = \mathbb{R}$ 时无须证明).

如果 $Ap = \lambda p$, 其中 $p \neq 0$, 则 $(Ap, p) = \lambda(p, p)$, 从而由 (a) 可得 $\lambda = \frac{(Ap,p)}{(p,p)} \in \mathbb{R}$. 由此可知若 A 是非负定的, 则 $\lambda \geqslant 0$; 若 A 是正定的, 则 $\lambda > 0$. (b) 得证.

如果 $Ap_1 = \lambda_1 p_1, Ap_2 = \lambda_2 p_2$, 则

$$(Ap_1, p_2) = \lambda_1(p_1, p_2) = (p_1, Ap_2) = \lambda_2(p_1, p_2).$$

所以 $(\lambda_1 - \lambda_2)(p_1, p_2) = 0$, 由此可知, 当 $\lambda_1 \neq \lambda_2$ 时 $(p_1, p_2) = 0$. 这就证得了 (c).

如果 $A \in \mathcal{L}(X)$, 由 Cauchy-Schwarz-Bunyakovskiĭ 不等式即得

$$\nu(A) := \sup_{x \neq 0} \frac{|(Ax, x)|}{\|x\|^2} \leqslant \|A\|.$$

为证得 (d), 只要证明 $\|A\| \leqslant \nu(A)$. 为此, 利用前述的算子范数 $\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}$ 在内积空间中也可表示为

$$\|A\| = \sup_{\substack{x \neq 0 \\ y \neq 0}} \frac{|(Ax, y)|}{\|x\|\|y\|} = \sup_{\|x\|=\|y\|=1} |(Ax, y)|$$

(定理 4.1-3). 令 x, y 满足 $(Ax, y) \neq 0$. $|(Ax, y)|$ 又可表示为

$$|(Ax, y)| = \frac{(Ax, y)\overline{(Ax, y)}}{|(Ax, y)|} = (A\tilde{x}, y),$$

其中 $\tilde{x} = \frac{\overline{(Ax,y)}}{|(Ax,y)|} x$. 由于 $(A\tilde{x}, y) = |(Ax, y)| \in \mathbb{R}$, 故有 $(A\tilde{x}, y) = (y, A\tilde{x})$; 又由于 A 自伴, 故有 $(y, A\tilde{x}) = (Ay, \tilde{x})$; 于是, $|(Ax, y)|$ 可进一步改写为

$$\begin{aligned}
|(Ax, y)| = (A\tilde{x}, y) &= \frac{1}{2}\{(A\tilde{x}, y) + (Ay, \tilde{x})\} \\
&= \frac{1}{4}\{(A(\tilde{x} + y), \tilde{x} + y) - (A(\tilde{x} - y), \tilde{x} - y)\}.
\end{aligned}$$

依次利用 $\nu(A)$ 的定义, 平行四边形法则和关系式 $\|\tilde{x}\| = \|x\|$, 便得

$$\begin{aligned}
|(Ax, y)| &\leqslant \frac{1}{4}\nu(A)\{\|\tilde{x} + y\|^2 + \|\tilde{x} - y\|^2\} \\
&= \frac{1}{2}\nu(A)\{\|\tilde{x}\|^2 + \|y\|^2\} = \frac{1}{2}\nu(A)\{\|x\|^2 + \|y\|^2\}.
\end{aligned}$$

因此,

$$\|A\| = \sup_{\|x\|=\|y\|=1} |(Ax, y)| \leqslant \nu(A),$$

这就是要证明的结果. □

4.11 紧自伴算子的谱定理

如所熟知, 任何 $n \times n$ 实对称阵或 $n \times n$ 复 Hermite 阵恰有 n 个实特征值 (按重数计) 它们可以用 Rayleigh 商的方法计算[27], 而且恰有 n 个相应的特征向量组成 $\mathbb{R}^n$ 或 $\mathbb{C}^n$ 的规范正交基. 值得注意的是在无限维内积空间中, 任何紧 (2.10 节) 的自伴 (4.10 节) 线性算子也具有相似的性质. 而且, 如果这个算子还是单射, 则它至多只有可数无限多个非零实特征值, 每一个均为有限重的 (即其对应的特征子空间均为有

[27] 见 Ciarlet [1987] 第 1.3 节.

限维的), 相应的特征向量组成极大规范正交系. 这就是下一个定理的核心, 即构成了这类算子的谱定理.

这个结果值得重视, 因为对这些算子特征值存在性的证明中并不需要借助于类似有限维情况下的行列式和特征多项式等概念.

注 相反地, 在"一般"的无限维赋范向量空间中, 对"一般"的线性算子的特征值和特征向量的研究则需要十分精细的工作[28) (如有限维情况所启示的, 相对于实对称或复 Hermite 矩阵的对角化定理, 还有 Jordan 标准形.)

定理 4.11-1 (具有无限维值域的紧自伴算子的谱定理) 设 $(X,(\cdot,\cdot))$ 为无限维内积空间, $A: X \to X$ 为紧的自伴线性算子, 具有无限维的值域, 则

(a) 存在 A 的特征值的无限序列 $(\lambda_n)_{n=1}^{\infty}$ 和相应的特征向量的无限序列 $(p_n)_{n=1}^{\infty}$, 满足

$$
\begin{aligned}
&|\lambda_1| = \|A\|, \quad |\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_n| \geqslant \cdots;\\
&\lambda_n \neq 0, \quad n \geqslant 1; \quad \lim_{n\to\infty} \lambda_n = 0;\\
&Ap_n = \lambda_n p_n, \quad n \geqslant 1,\\
&(p_k, p_\ell) = \delta_{k,\ell}, \quad k,\ell \geqslant 1,
\end{aligned}
$$

$$
\begin{aligned}
|\lambda_1| &= \frac{|(Ap_1,p_1)|}{\|p_1\|^2} = \sup_{x\neq 0} \frac{|(Ax,x)|}{\|x\|^2},\\
|\lambda_n| &= \frac{|(Ap_n,p_n)|}{\|p_n\|^2} = \sup_{\substack{x\neq 0\\ (x,p_k)=0,1\leqslant k\leqslant n-1}} \frac{|(Ax,x)|}{\|x\|^2}, \quad n \geqslant 2.
\end{aligned}
$$

(b) 对任何向量 $x \in X$,

$$
Ax = \sum_{n=1}^{\infty} \lambda_n (x,p_n) p_n.
$$

(c) 设 λ 为 A 的任何非零特征值, 则存在 $n \geqslant 1$, 使得 $\lambda_n = \lambda$. 而且, 集合 $I(\lambda) := \{n \geqslant 1; \lambda_n = \lambda\}$ 为有限集,

$$
\{p \in X; Ap = \lambda p\} = \mathrm{Span}(p_n)_{n\in I(\lambda)}.
$$

(d) A 的核空间为

$$
\mathrm{Ker}\, A = (\mathrm{Span}(p_n)_{n=1}^{\infty})^{\perp}.
$$

证明 为方便起见, 把证明分为 (i) 到 (vii) 几部分. 如前已证, 自伴算子的所有特征值均为实数 (定理 4.10-1(b)).

[28) 简短的介绍可见 Taylor [1958] 的第 5 章或 Taylor & Lay [1980] 的第 5 章. 详尽的论述见 Dunford & Schwartz [1963].

(i) 存在特征值 λ_1, 和相应的特征向量 p_1 满足

$$Ap_1 = \lambda_1 p_1, \quad \|p_1\| = 1,$$
$$0 < |\lambda_1| = \|A\| = \sup_{x \neq 0} \frac{|(Ax, x)|}{\|x\|^2} = \frac{|(Ap_1, p_1)|}{\|p_1\|^2}.$$

由 A 的自伴性可得 (定理 4.10-1(d)) $\sup_{\|x\|=1} |(Ax, x)| = \|A\|$, 又由 $A \neq 0$ (由假设, 值域 $A(X)$ 为无限维的) 得 $\|A\| > 0$. 因此, 存在单位向量的序列 $(x_n)_{n=1}^{\infty}$ 使得对所有 $n \geqslant 1, \|x_n\| = 1$, 当 $n \to \infty$ 时 $|(Ax_n, x_n)| \to \|A\|$. 从而, 存在子列 $(x_m)_{m=1}^{\infty}$ 和 $\lambda_1 \in \mathbb{R}$, 使得

$$\|x_m\| = 1, \quad m \geqslant 1; (Ax_m, x_m) \underset{m\to\infty}{\longrightarrow} \lambda_1,$$
$$|\lambda_1| = \sup_{\|x\|=1} |(Ax, x)| = \|A\| > 0.$$

由于 $(x_m)_{m=1}^{\infty}$ 是有界序列, 又由假设, A 是紧的, 从而存在序列 $(x_m)_{m=1}^{\infty}$ 的子列 $(x_\ell)_{\ell=1}^{\infty}$, 使得序列 $(Ax_\ell)_{\ell=1}^{\infty}$ 在 X 中收敛. 注意到

$$\begin{aligned}\|Ax_\ell - \lambda_1 x_\ell\|^2 &= \|Ax_\ell\|^2 - 2\lambda_1(Ax_\ell, x_\ell) + \lambda_1^2 \\ &\leqslant \|A\|^2 - 2\lambda_1(Ax_\ell, x_\ell) + \lambda_1^2\end{aligned}$$

(即使在复的情况下, $(Ax_\ell, x_\ell) \in \mathbb{R}$; 见定理 4.10-1(a)), 又

$$(\|A\|^2 - 2\lambda_1(Ax_\ell, x_\ell) + \lambda_1^2) \underset{\ell\to\infty}{\longrightarrow} 0,$$

可得

$$(Ax_\ell - \lambda_1 x_\ell) \underset{\ell\to\infty}{\longrightarrow} 0,$$

从而

$$x_\ell = \left\{-\frac{1}{\lambda_1}(Ax_\ell - \lambda_1 x_\ell) + \frac{1}{\lambda_1} Ax_\ell\right\} \underset{\ell\to\infty}{\longrightarrow} p_1 := \frac{1}{\lambda_1} \lim_{\ell\to\infty} Ax_\ell.$$

又由于对所有的 $\ell \geqslant 1$, $\|x_\ell\| = 1$, 所以

$$\|p_1\| = 1.$$

最后, 由 A 的连续性

$$Ap_1 = A(\lim_{\ell\to\infty} x_\ell) = \lim_{\ell\to\infty} Ax_\ell = \lambda_1 p_1,$$

所以 λ_1 满足或者 $\lambda_1 = \|A\|$ 或者 $\lambda_1 = -\|A\|$, 它是 A 的一个特征值.

(ii) 存在特征值的无限序列 $(\lambda_n)_{n=1}^{\infty}$ 和特征向量的无限序列 $(p_n)_{n=1}^{\infty}$, 同时满足

$$Ap_n = \lambda_n p_n, \quad n \geqslant 1; \quad (p_k, p_\ell) = \delta_{k\ell}, \quad k, \ell \geqslant 1;$$

且对所有的 $n \geqslant 2$,

$$0 < |\lambda_n| = \sup_{\substack{x \neq 0 \\ (x,p_k)=0, 1 \leqslant k \leqslant n-1}} \frac{|(Ax,x)|}{\|x\|^2} = \frac{|(Ap_n,p_n)|}{\|p_n\|^2} \leqslant |\lambda_{n-1}| \leqslant \cdots \leqslant |\lambda_1|.$$

定义子空间

$$X_2 := \{x \in X; (x,p_1) = 0\},$$

其中 p_1 是 (i) 中出现的特征向量. 因为对任何 $x \in X_2$,

$$(Ax,p_1) = (x,Ap_1) = \lambda_1(x,p_1) = 0,$$

所以 X_2 在 A 作用下的直接像 $A(X_2)$ 包含于 X_2.

显然, A 在 X_2 上的限制 A_2 也是紧的自伴线性算子. 而且 $A_2 \neq 0$; 这是因为如果 $A_2 = 0$, 则对所有的 $x \in X$, 因为 $(x-(x,p_1)p_1) \in X_2$, 故而

$$A(x-(x,p_1)p_1) = 0,$$

所以对所有的 $x \in X$

$$Ax = \lambda_1(x,p_1)p_1,$$

这与 $A(X)$ 为无限维的相矛盾. 于是, 可以将 (i) 的论证用于 $A_2 : X_2 \to X_2$, 便得存在 $\lambda_2 \in \mathbb{R}$ 和向量 $p_2 \in X_2$, 同时满足

$$A_2p_2 = Ap_2 = \lambda_2 p_2, \quad (p_2,p_1) = 0, \quad \|p_2\| = 1,$$

$$0 < |\lambda_2| = \|A_2\| = \sup_{\substack{x \neq 0 \\ x \in X_2}} \frac{|(A_2x,x)|}{\|x\|^2} = \sup_{\substack{x \neq 0 \\ x \in X_2}} \frac{|(Ax,x)|}{\|x\|^2}$$

$$\leqslant \sup_{x \neq 0} \frac{|(Ax,x)|}{\|x\|^2} = |\lambda_1|.$$

重复以上过程: 假设对某个 $n \geqslant 2$, 已经取到特征值 $\lambda_k, 1 \leqslant k \leqslant n$, 和相应的特征向量 $p_k, 1 \leqslant k \leqslant n$, 同时满足

$$Ap_k = \lambda_k p_k, \quad (p_k,p_\ell) = \delta_{k\ell}, \quad 1 \leqslant k,\ \ell \leqslant n,$$

$$0 < |\lambda_n| = \sup_{\substack{x \neq 0 \\ (x,p_k)=0, 1 \leqslant k \leqslant n-1}} \frac{|(Ax,x)|}{\|x\|^2} \leqslant |\lambda_{n-1}| \leqslant \cdots \leqslant |\lambda_1|.$$

定义子空间

$$X_{n+1} := \{x \in X; (x,p_k) = 0, 1 \leqslant k \leqslant n\}.$$

因为对所有的 $x \in X_{n+1}$ 和 $1 \leqslant k \leqslant n$,

$$(Ax,p_k) = (x,Ap_k) = \lambda_k(x,p_k) = 0,$$

所以 X_{n+1} 在 A 作用下的直接像 $A(X_{n+1})$ 包含于 X_{n+1} 中.

显然, A 在 X_{n+1} 上的限制 A_{n+1} 仍是紧的自伴线性算子. 而且 $A_{n+1} \neq 0$, 这是因为如果 $A_{n+1} = 0$, 则对所有的 $x \in X$, 由于 $(x - \sum_{k=1}^{n}(x, p_k)p_k) \in X_{n+1}$, 故而

$$A\left(x - \sum_{k=1}^{n}(x, p_k)p_k\right) = 0,$$

所以对所有的 $x \in X$

$$Ax = \sum_{k=1}^{n} \lambda_k(x, p_k)p_k,$$

这与 $A(X)$ 是无限维空间相矛盾. 再将 (i) 的论证用于 A_{n+1}, 可知存在 $\lambda_{n+1} \in \mathbb{R}$ 和 $p_{n+1} \in X_{n+1}$, 同时满足

$$\begin{aligned}
&A_{n+1}p_{n+1} = Ap_{n+1} = \lambda_{n+1}p_{n+1};\\
&(p_{n+1}, p_k) = 0, \quad 1 \leqslant k \leqslant n; \quad \|p_{n+1}\| = 1;\\
&0 < |\lambda_{n+1}| = \sup_{\substack{x \neq 0\\ x \in X_{n+1}}} \frac{|(A_{n+1}x, x)|}{\|x\|^2} = \sup_{\substack{x \neq 0\\ (x, p_k) = 0, 1 \leqslant k \leqslant n}} \frac{|(Ax, x)|}{\|x\|^2}\\
&\qquad\qquad \leqslant \sup_{\substack{x \neq 0\\ (x, p_k) = 0, 1 \leqslant k \leqslant n-1}} \frac{|(Ax, x)|}{\|x\|^2} = |\lambda_n| \leqslant \cdots \leqslant |\lambda_1|.
\end{aligned}$$

于是, 对 $n \geqslant 2$, 所述的性质成立.

(iii) 在 (ii) 中出现的特征值 $\lambda_n, n \geqslant 1$, 满足 $\lim_{n\to\infty} \lambda_n = 0$.

若不然, 则存在 $\delta > 0$, 使得对所有的 $n \geqslant 1$, 有

$$|\lambda_n| \geqslant \delta > 0$$

(因为 $(|\lambda_n|_{n=1}^{\infty})$ 是单调下降的序列, 见 (ii)). 于是, 序列 $\left(\frac{1}{\lambda_n}p_n\right)_{n=1}^{\infty}$ 在 X 中有界, 由 A 的紧性 (定理 2.10-1(b)), 存在子列 $\left(\frac{1}{\lambda_{\sigma(n)}}p_{\sigma(n)}\right)_{n=1}^{\infty}$ 使得序列

$$\left(A\left(\frac{1}{\lambda_{\sigma(n)}}p_{\sigma(n)}\right)\right)_{n=1}^{\infty} = (p_{\sigma(n)})_{n=1}^{\infty}$$

收敛. 但这是不可能的, 因为在 (ii) 中得到的特征向量是规范正交的, 所以对任何 $k \neq \ell$, 均有

$$\|p_k - p_\ell\|^2 = \|p_k\|^2 + \|p_\ell\|^2 = 2.$$

这就证得定理 4.11-1 中 (a) 的所有性质.

(iv) 对任何向量 $x \in X$, 向量 $Ax \in X$ 可表示为收敛级数的和: $Ax = \sum_{k=1}^{\infty} \lambda_k(x, p_k)p_k$.

对任意给定的 $x \in X$ 和任意整数 $n \geqslant 1$, 定义向量

$$x_n := x - \sum_{k=1}^{n} (x, p_k) p_k,$$

由 (ii) 已见它属于子空间

$$X_{n+1} = \{x \in X; (x, p_k) = 0, 1 \leqslant k \leqslant n\} = (\mathrm{Span}(p_k)_{k=1}^n)^{\perp}.$$

因为对每个 $n \geqslant 1$, 向量 x_n 正交于向量 $\sum_{k=1}^n (x, p_k) p_k$, 由 Pythagoras 定理可得

$$\|x_n\| \leqslant \|x\|.$$

因为 $Ax_n = A_{n+1} x_n$ (原因是 $x_n \in X_{n+1}, A_{n+1} = A|_{X_{n+1}}$), 又由 (ii)

$$\|A_{n+1}\| = \sup_{\substack{x \neq 0 \\ x \in X_{n+1}}} \frac{|(A_{n+1} x, x)|}{\|x\|^2} = |\lambda_{n+1}|,$$

所以

$$\begin{aligned} \left\| Ax - \sum_{k=1}^{n} \lambda_k (x, p_k) p_k \right\| &= \|Ax_n\| = \|A_{n+1} x_n\| \\ &\leqslant \|A_{n+1}\| \|x_n\| = |\lambda_{n+1}| \|x_n\| \leqslant |\lambda_{n+1}| \|x\|. \end{aligned}$$

由 (iii), $\lim_{n \to \infty} |\lambda_{n+1}| = 0$, 所以 $\lim_{n \to \infty} \{Ax - \sum_{k=1}^n \lambda_k (x, p_k) p_k\} = 0$. 这就证得了 (b).

(v) 在 (ii) 中通过递推过程已经取到了 A 的全部非零特征值.

设 $\lambda \neq 0, p \neq 0$ 满足 $Ap = \lambda p$, 于是 $Ap \neq 0$. 如果对任何 $k \geqslant 1$ 均有 $\lambda \neq \lambda_k$, 则对所有的 $k \geqslant 1$, 均有 $(p, p_k) = 0$ (定理 4.10-1(c)). 因此, 由 (iv) 可得

$$Ap = \lim_{n \to \infty} \left(\sum_{k=1}^{n} \lambda_k (p, p_k) p_k \right) = 0,$$

此为矛盾.

(vi) 任意给定 A 的非零特征值 λ, 定义集合

$$I(\lambda) := \{n \geqslant 1; \lambda_n = \lambda\},$$

它是非空的 (由 (v) 可知), 且是有限集 (由 (iii), $\lim_{n \to \infty} \lambda_n = 0$). 则有

$$\{p \in X; Ap = \lambda p\} = \mathrm{Span}(p_n)_{n \in I(\lambda)}.$$

换言之, A 的每个非零特征值均为有限重数的, A 的所有对应于非零特征值的特征子空间均已在 (ii) 的递推过程中取到.

由 (ii), 存在特征向量 $p_n, n \in I(\lambda)$, 使得

$$Ap_n = \lambda p_n,$$
$$(p_k, p_\ell) = \delta_{k\ell}, \quad k, \ell \in I(\lambda).$$

因此 $\text{Span}(p_n)_{n\in I(\lambda)} \subset \{p \in X; Ap = \lambda p\}$. 为证明这个包含关系实际上是个等式, 用反证法. 如果存在向量 $\tilde{p} \in X$, 使得 $A\tilde{p} = \lambda\tilde{p}, \tilde{p} \neq 0$, 且 $\tilde{p} \bar{\in} \text{Span}(p_n)_{n\in I(\lambda)}$. 将 Gram-Schmidt 规范正交化方法 (定理 4.8-1) 应用于向量 $\tilde{p}$ 和 p_n, $n \in I(\lambda)$, 则可得向量 $p \in \text{Span}\{(p_n)_{n\in I(\lambda)}, \tilde{p}\}$ 满足 $\|p\| = 1$, 且对所有的 $n \in I(\lambda)$,

$$(p, p_n) = 0.$$

因为 p 是向量 $\tilde{p}$ 和 $p_n, n \in I(\lambda)$, 的线性组合, 所以

$$Ap = \lambda p.$$

这样, 由定理 4.10-1(c), 因为对任何 $n \bar{\in} I(\lambda), \lambda \neq \lambda_n$, 所以

$$(p, p_n) = 0, \quad n \bar{\in} I(\lambda),$$

因此, 对任何 $n \geqslant 1$, 非零向量 p 满足 $(p, p_n) = 0$. 用与 (v) 同样的论证即可导致矛盾.

这就证得了 (c) 中所述的全部性质.

(vii) 余下只要证明

$$\text{Ker } A = (\text{Span}(p_n)_{n=1}^{\infty})^{\perp}$$

(正交补定义见第 4.5 节). 设 $x \in X$ 满足 $Ax = 0$. 于是, 对任何 $n \geqslant 1$,

$$(x, p_n) = \frac{1}{\lambda_n}(x, \lambda_n p_n) = \frac{1}{\lambda_n}(x, Ap_n) = \frac{1}{\lambda_n}(Ax, p_n) = 0,$$

这表明 $x \in (\text{Span}(p_n)_{n=1}^{\infty})^{\perp}$.

反之, 若 $x \in (\text{Span}(p_n)_{n=1}^{\infty})^{\perp}$, 则对所有的 $n \geqslant 1$, 有 $(x, p_n) = 0$, 故而由 (iv) 可得 $Ax = 0$. (d) 得证. □

当然, 如果 A 是非负定的或正定的, 则在定理 4.11-1 中取得的特征值 $\lambda_n, n \geqslant 1$, 满足

$$\lambda_n > 0, \quad n \geqslant 1; \quad 且 \quad \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n \cdots,$$

它们的特征不再需要用绝对值来表示, 其形式为

$$\lambda_1 = \sup_{x\neq 0} \frac{(Ax, x)}{\|x\|^2}, \quad \lambda_n = \sup_{\substack{x\neq 0 \\ (x, p_k)=0, 1\leqslant k\leqslant n-1}} \frac{(Ax, x)}{\|x\|^2}.$$

注　定理 4.11-1 的逆也成立, 见习题 4.11-1.

为完整起见, 我们再来考察更简单的情况, 即 $A: X \to X$ 是值域为有限维空间的连续自伴线性算子 (此时 A 必是紧的, 见定理 2.10-1(d)). 这类算子包含了有限维空间中可表示为实对称或复 Hermite 矩阵的算子.

定理 4.11-2 (具有有限维值域的连续自伴线性算子的谱定理) 设 $(X, (\cdot,\cdot))$ 为内积空间, $A: X \to X$ 为连续自伴线性算子, 其值域为有限维 $N \geqslant 1$. 则

(a) 恰存在 N 个 A 的非零特征值 λ_n 和相应的特征向量 $p_n, 1 \leqslant n \leqslant N$, 满足

$$|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots \geqslant |\lambda_N| > 0,$$

而且

$$\begin{aligned} &Ap_n = \lambda_n p_n, \quad 1 \leqslant n \leqslant N; \quad (p_k, p_\ell) = \delta_{k\ell}, \quad 1 \leqslant k,\ \ell \leqslant N, \\ &|\lambda_1| = \sup_{x \neq 0} \frac{|(Ax,x)|}{\|x\|^2}, \end{aligned}$$

如果 $N \geqslant 2$, 则对 $2 \leqslant n \leqslant N$,

$$|\lambda_n| = \sup_{\substack{x \neq 0 \\ (x,p_k)=0, 1 \leqslant k \leqslant n-1}} \frac{|(Ax,x)|}{\|x\|^2}.$$

(b) 设 λ 为 A 的任一非零特征值, 则存在 $n \in \{1, \cdots, N\}$, 使得 $\lambda_n = \lambda$, 而且

$$\{p \in X; Ap = \lambda p\} = \mathrm{Span}(p_n)_{n \in I(\lambda)},$$

其中 $I(\lambda) := \{n \in \{1, \cdots, N\}; \lambda_n = \lambda\}$.

(c) 对任何向量 $x \in X$,

$$Ax = \sum_{n=1}^{N} \lambda_n (x, p_n) p_n.$$

(d) A 的核空间可表示为

$$\mathrm{Ker}\ A = (\mathrm{Span}(p_n)_{n=1}^{N})^{\perp}.$$

证明　这里的证明只需对定理 4.11-2 证明中从 (i) 到 (vii) 作简单的修改, 下面依然使用那里的记号.

因为 A 是紧的, 且 $A \neq 0$ (由假设 A 的值域维数 $N \geqslant 1$), 逐句重复即知 (i) 成立.

如果 $N = 1$, A 的值域就与子空间 $\mathrm{Span}(p_1)$ 一致. 如果 $N \geqslant 2$, 则 $A_2 \neq 0$, 从而可以开始 (ii) 的递推过程. 本质的区别在于这个递推过程在第 N 步后必定终止, 这是因为 $A_{N+1} = 0$, 其中 A_{N+1} 表示 A 在子空间 X_{N+1} 上的限制, 其中

$$X_{N+1} = \{x \in X; (x, p_k) = 0, 1 \leqslant k \leqslant N\} = (\mathrm{Span}(p_k)_{k=1}^{N})^{\perp}.$$

注意到 $A_{N+1}=0$, 即得对所有的 $x\in X$,

$$Ax=\sum_{k=1}^{N}\lambda_k(x,p_k)p_k,$$

这就证明了 A 的值域是 N 维的 (特征值 $\lambda_k, 1\leqslant k\leqslant N$ 均为非零的; 特征向量 $p_k, 1\leqslant k\leqslant N$, 线性无关; $Ap_k=\lambda_k p_k, 1\leqslant k\leqslant N$). 这就证得 (a) 和 (c).

从 (v) 到 (vii) 的论证可以几乎逐字重复, 这样便证得 (b) 与 (d). □

在结束本节的分析时, 我们来说明定理 4.11-1 提供了一个简单的推论, 即当算子 A 还是单射时, 它给出一个构造极大规范正交系和 Hilbert 基的重要手段.

定理 4.11-3 (a) 设 $(X,(\cdot,\cdot))$ 为无限维内积空间, $A:X\to X$ 为紧的自伴线性算子, 而且是单射. 则定理 4.11-1 中的特征向量族 $(p_n)_{n=1}^{\infty}$ 是 X 中的极大规范正交系.

(b) 如果 X 还是可分的 Hilbert 空间, 则定理 4.11-1 中的特征向量族 $(p_n)_{n=1}^{\infty}$ 是 X 的一个 Hilbert 基.

证明 首先, 注意到因为 A 是单射, 且 X 为无限维空间, 所以 A 的值域必是无限维的. 因此, 定理 4.11-1 的所有假设均满足.

由假设 $\operatorname{Ker} A=\{0\}$, 所以

$$(\operatorname{Span}(p_n)_{n=1}^{\infty})^{\perp}=\{0\},$$

即 $(p_n)_{n=1}^{\infty}$ 是 X 中的极大规范正交系, 或有当 X 为可分 Hilbert 空间时, 它是一个 Hilbert 基. □

在定理 4.11-3(b) 的假设下, 由定理 4.9-1 和 4.11-1, 两个值得注意的公式同时成立, 即对任何 $x\in X$.

$$x=\sum_{n=1}^{\infty}(x,p_n)p_n,\quad Ax=\sum_{n=1}^{\infty}\lambda_n(x,p_n)p_n.$$

习题

4.11-1 设 X 为可分 Hilbert 空间, $(e_n)_{n=1}^{\infty}$ 为 X 的 Hilbert 基 (4.9 节), $(\lambda_n)_{n=1}^{\infty}$ 为有界实数列.

(1) 证明: 对任何 $x\in X$, 级数 $\sum_{n=1}^{\infty}\lambda_n(x,e_n)e_n$ 在 X 中收敛.

(2) 对任何 $x\in X$, 令 $Ax:=\sum_{n=1}^{\infty}\lambda_n(x,e_n)e_n$. 证明这样定义的映射 $A:X\to X$ 是连续的自伴线性算子.

(3) 证明对每个 $n\geqslant 1$, λ_n 是 A 的特征值, e_n 是相应的特征向量.

(4) 证明: 如果对所有的 $n\geqslant 1, \lambda_n\neq 0$, 则算子 A 为单射.

(5) 证明: 如果 $\lim_{n\to\infty}\lambda_n=0$, 则算子 A 是紧的.

第 5 章　线性泛函分析中的重要定理

引言

本章着重于给出大部分线性泛函分析主要定理的证明. 这些定理的共同特点是, 它们有赖于下述两个基本结果或其中的某一个: Baire 定理 (定理 5.1-2) 和赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1).

Baire 定理断言, Banach 空间 (或更一般地, 完备距离空间) 中可列无穷多个稠密开子集的交仍为稠密的.

Baire 定理的直接结果包括多项式空间的不完备性 (不管装备何种范数, 定理 5.1-4) 及存在 "许多" 处处不可微的连续函数 (定理 5.2-1).

Baire 定理的另一个结果是 Banach-Steinhaus 定理, 也称为一致有界原理 (定理 5.3-1), 它是线性泛函分析的基石之一. 例如, 由这个定理可以推得, 存在连续函数, 其 Lagrange 多项式插值不一致收敛 (定理 5.4-2); 此外还存在连续函数, 其 Fourier 级数不一致收敛 (定理 5.5-1).

线性泛函分析的另外两个基石是 Banach 开映射定理 (定理 5.6-1) 和 Banach 闭图像定理 (定理 5.7-1), 它们也是 Baire 定理的推论. 我们用两个重要的应用来说明它们的效力, 第一个是关于在最少假设下, 一个微分算子之逆的连续性 (定理 5.6-3); 而另一个是令人难以想象的 Hellinger-Toeplitz 定理 (定理 5.7-2), 它断定 Hilbert 空间中任一自伴算子都自动地连续.

Hahn-Banach 定理 (定理 5.9-1) 则属不同的情况, 其证明需要选择公理. 这个定理断定, 在任何赋范向量空间 X 中, X 的任一子空间上的任一连续线性形式都可以保持原来同样的范数延拓为全空间 X 上的连续线性形式.

由这一定理推得的结果也给人留下深刻的印象: 这其中包括线性泛函分析中的一

些基本定理. 例如, Hahn-Banach 定理的 “几何形式” (定理 5.10-1 和定理 5.10-2), 它阐明了在任一赋范向量空间中, 如何去 “分离凸集”; 还有意义深刻的 Banach 闭值域定理 (定理 5.11-5 和定理 5.11-6), 它会同 Banach 开映射定理, 给出了利用对偶算子判定一个线性算子是否是满射的一个非常简单的充分条件, 或更一般地, 利用对偶算子给出其映像的刻画.

值得注意的是, 在一定意义上, 利用 Hahn-Banach 定理可以将在 Hilbert 空间成立的性质推广到任意的赋范向量空间 X. 例如, 对偶算子的概念 (5.11 节), 它推广伴随算子的概念; 还有, 到一个闭子空间上的投影定理 (定理 5.9-7) 等.

我们以两个基本概念: 弱收敛 (5.12 节) 和自反 Banach 空间, 结束本章. 对它们的分析通常既有赖于 Baire 定理也依赖于 Hahn-Banach 定理. 这一系列的分析, 最后以两个主要结果: Banach-Saks-Mazur 定理 (定理 5.13-1) 和 Banach-Eberlein-Šmulian 定理 (5.14-4) 结束. 它们在后续讨论中将起重要的作用.

说起来似乎不太合理, 本章中确定的线性泛函分析中的 “主要定理”, 在专门讨论线性偏微分方程的下一章中并没有用到多少; 但相比之下, 在最后一章中它们却连同非线性泛函分析的基本定理被多次用于确立非线性偏微分方程解的存在性.

5.1 Baire 定理; 首选应用: 多项式空间的不完备性

Baire 定理 (定理 5.1-2) 是线性泛函分析中的两个核心定理之一, 另一个是 Hahn-Banach 定理 (5.8—5.10 节). 由 Baire 定理得到的意义深远的结论中包括这样一些基本定理, 如 Banach-Steinhaus 定理 (定理 5.3-1), Banach 开映射定理 (定理 5.6-1), 以及 Banach 闭图像定理 (定理 5.7-1) 等.

虽然 Baire 定理在本章的其余部分只应用于 Banach 空间, 但我们将在更一般的完备距离空间的框架下给出该定理的证明, 因为这样的一般性并不会增加额外的工作量.

Baire 定理基于完备距离空间的下述重要性质 (下面 $\operatorname{diam} A = \sup\{d(x,y); x \in A, y \in A\} \in [0,\infty]$ 表示距离空间 (X,d) 的一子集 A 的直径; 见 1.10 节).

定理 5.1-1 (Cantor 交集定理) 设 X 是完备距离空间, 而 $(A_n)_{n=0}^{\infty}$ 为 X 中满足如下性质的非空闭子集 A_n 组成的序列, 它们满足

$$A_0 \supset A_1 \supset \cdots \supset A_n \supset A_{n-1} \supset \cdots \text{ 且}$$
$$\text{当 } n \to \infty \text{ 时,}\quad \operatorname{diam} A_n \to 0.$$

则存在 $x \in X$ 使

$$\bigcap_{n=0}^{\infty} A_n = \{x\}.$$

证明 对每个 $n \geqslant 0$, 取一个元素 $x_n \in A_n$ (每一子集 A_n 都是非空的). 因为 $A_m \subset A_n$ 对所有 $m \geqslant n$ 成立, 这意味着

$$d(x_m, x_n) \leqslant \operatorname{diam} A_n, \quad \forall\, m \geqslant n,$$

而且当 $n \to \infty$ 时, $\operatorname{diam} A_n \to 0$, 所以 $(x_n)_{n=0}^{\infty}$ 是 Cauchy 序列. 设 $x := \lim_{n\to\infty} x_n$.

给定任一整数 $n \geqslant 0$, 对所有 $m \geqslant n$, 有 $x_m \in A_n$. 又因 A_n 是闭集, 所以 $x = \lim_{m\to\infty} x_m \in A_n$. 这就得到 $x \in \bigcap_{n=0}^{\infty} A_n$.

假定交集 $\bigcap_{n=0}^{\infty} A_n$ 包含另一点 $y \neq x$. 则存在 $n_0 > 0$ 使得 $\operatorname{diam} A_{n_0} < d(x, y)$. 但是因为 x 与 y 均属于 A_{n_0}, 所以 $d(x, y) \leqslant \operatorname{diam} A_{n_0}$, 与前述矛盾. 故 $\bigcap_{n=0}^{\infty} A_n = \{x\}$. □

注 (1) 当 $n \to \infty$ 时 $\operatorname{diam} A_n \to 0$ 这一假设是实质性的. 作为例子可以考察特殊情况: $X = \mathbb{R}$ 而 $A_n = [n, \infty)$, $n \geqslant 1$.

(2) 定理 5.1-1 中所确立的性质实际上刻画了所讨论的距离空间的完备性 (习题 5.1-1).

(3) Cantor 交集定理将在以后确立 Ekeland 变分原理 (定理 9.8-1) 中起着实质性的作用 (习题 9.8-1).

定理 5.1-2 (Baire 定理[1]) 设 X 是完备距离空间. 则下述两个等价性质成立:

(a) 设 $(F_n)_{n=0}^{\infty}$ 是 X 中的闭子集序列, 且对所有 $n \geqslant 0$, $\operatorname{int} F_n = \varnothing$. 则

$$\operatorname{int}\left(\bigcup_{n=0}^{\infty} F_n\right) = \varnothing.$$

(b) 设 $(O_n)_{n=0}^{\infty}$ 是 X 中的开子集序列, 且对所有 $n \geqslant 0$, $\overline{O}_n = X$. 则

$$\overline{\bigcap_{n=0}^{\infty} O_n} = X.$$

证明 为印刷方便起见, 我们采用符号 $\operatorname{int} A$ 而不用 $\mathring{A}$.

(i) 我们从证明性质 (a) 与 (b) 等价开始. 先假定性质 (a) 成立.

首先我们注意, 对任意的 $A \subset X$, $\overline{A} = X - \operatorname{int}(X - A)$ 意味着

$$\overline{A} = X \quad \text{当且仅当} \quad \operatorname{int}(X - A) = \varnothing.$$

现给定开集 $O_n \subset X$, $n \geqslant 0$, 使得对所有 $n \geqslant 0$ 成立 $\overline{O}_n = X$. 设 $F_n := X - O_n$. 则闭集 F_n 就满足: 对所有 $n \geqslant 0$, $\operatorname{int} F_n = \varnothing$. 所以由 (a) 得 $\operatorname{int}(\bigcup_{n=0}^{\infty} F_n) = \varnothing$. 但

[1] 这个定理首先是对 $X = [a, b]$ 这一特殊情况给出证明. 见:

R. Baire [1899]: Sur les fonctions de variables réelles. Annali di Matematica Pura ed Applicata 3, 1–123.

由 de Morgan 定律 (1.3 节),

$$\bigcup_{n=0}^{\infty} F_n = \bigcup_{n=0}^{\infty}(X - O_n) = X - \bigcap_{n=0}^{\infty} O_n.$$

因此 $\operatorname{int}(X - \bigcap_{n=0}^{\infty} O_n) = \varnothing$, 这就有 $\overline{\bigcap_{n=0}^{\infty} O_n} = X$. 故性质 (b) 成立.

由性质 (b) 推得 (a) 的证明是类似的. 此时要注意到对任意的 $B \subset X$, 关系式 $\overline{X - B} = X - \operatorname{int} B$ 成立, 意味着

$$\operatorname{int} B = \varnothing \quad \text{当且仅当} \quad \overline{X - B} = X.$$

(ii) 我们现在证明 (a). 首先注意到, X 的一子集 A 有非空内部的充分必要条件是存在一个包含于 A 中的非空开子集 $O \subset X$ 或等价地 $O \cap (X - A) = \varnothing$. 这样就得

$$\operatorname{int} A = \varnothing \text{ 的充分必要条件是对所有的非空开子集 } O \subset X,$$
$$O \cap (X - A) \neq \varnothing.$$

现设 X 是完备距离空间, 而 F_n $(n \neq 0)$ 是 X 的闭子集使得对所有 $n \geqslant 0$ 有 $\operatorname{int} F_n = \varnothing$. 我们希望证明 $\operatorname{int} \bigcup_{n=0}^{\infty} F_n = \varnothing$ 或等价地证明, 对给定的任意非空开子集 $O \subset X$, 均有

$$O \cap \left(X - \bigcup_{n=0}^{\infty} F_n\right) \neq \varnothing.$$

给定一个非空开集 $O \subset X$, 置 $O_0 := O$. 由于 $\operatorname{int} F_0 = \varnothing$ 而 O_0 是开集, 故 $O_0 \cap (X - F_0)$ 是 X 的一个非空开子集. 所以存在一个非空开子集 $O_1 \subset X$ 使得

$$\overline{O}_1 \subset O_0 \cap (X - F_0) \quad \text{且} \quad \operatorname{diam} \overline{O}_1 < 1$$

(如以 $O_0 \cap (X - F_0)$ 中任一点为球心、半径充分小的球 O_1 即满足这要求). 由于 $\operatorname{int} F_1 = \varnothing$ 且 O_1 是开集, 故 $O_1 \cap (X - F_1)$ 是 X 的非空开子集, 同样地存在非空开子集 $O_2 \subset X$ 使得

$$\overline{O}_2 \subset O_1 \cap (X - F_1) \quad \text{且} \quad \operatorname{diam} \overline{O}_2 < 1/2.$$

以此方式继续, 我们可以构造一个非空开子集序列 $O_n \subset X$, $n \geqslant 0$, 使得

$$\overline{O}_{n+1} \subset O_n \cap (X - F_n) \quad \text{且} \quad \operatorname{diam} \overline{O}_{n+1} < \frac{1}{n+1}, \quad n \geqslant 0.$$

非空闭子集 $\overline{O}_n (n \geqslant 0)$ 显然满足 Cantor 交集定理 (定理 5.1-1) 的所有假设. 所以存在 $x \in X$ 使得 $\{x\} = \bigcap_{n=0}^{\infty} \overline{O}_n$.

一方面, $x \in \overline{O}_1 \subset O_0 = O$, 即 $x \in O$. 而另一方面, 关系式 $x \in \overline{O}_{n+1}$ 且 $\overline{O}_{n+1} \subset X - F_n$ 对所有 $n \geqslant 0$ 成立意味着 $x \in \bigcap_{n=0}^{\infty}(X - F_n)$. 注意到 $\bigcap_{n=0}^{\infty}(X - F_n) =$

$X - \bigcup_{n=0}^{\infty} F_n$, 我们有

$$x \in O \cap \left(X - \bigcup_{n=0}^{\infty} F_n \right).$$

这就说明, 集合 $O \cap (X - \bigcup_{n=0}^{\infty} F_n)$ 非空, 定理证毕. □

在应用中, Baire 定理常采取下述定理中性质 (a) 或 (b) 的形式 (这两条性质可立即由定理 5.1-2(a) 得出).

定理 5.1-3 设 X 是距离空间, 令 F_n $(n \geqslant 0)$ 是 X 的闭子集使得 $X = \bigcup_{n=0}^{\infty} F_n$.

(a) 如果对所有 $n \geqslant 0$, 有 $\operatorname{int} F_n = \varnothing$, 则 X 是不完备的.

(b) 如果 X 是完备的, 则存在 $n_0 \geqslant 0$ 使得 $\operatorname{int} F_{n_0} \neq \varnothing$.

作为例子, 考察所有有理数组成的可列无限集 $\mathcal{Q} = \bigcup_{n=0}^{\infty}\{q_n\}$ (以任意的方式排序), 在其中对所有 $p, q \in \mathcal{Q}$ 装备以由 $d(p,q) = |p-q|$ 定义的通常的距离 d. 因为 $\mathcal{Q}$ 的每一子集 $\{q_n\}$ 都是闭的而且内部是空集, 由定理 5.1-3(a) 得, $(\mathcal{Q}, d)$ 是不完备的 (这一结论当然也可直接得出: 例如, 如下定义有理数列 $(r_n)_{n=0}^{\infty}$, $r_0 = 0$ 而 $r_n = \dfrac{1}{4} + r_{r-1} - \dfrac{1}{2}(r_{n-1})^2$, $n \geqslant 1$. 它是 Cauchy 序列但在 $\mathcal{Q}$ 中不收敛).

同样可以立即得到定理 5.1-3(b) 如下值得关注的推论: 平面 $\mathbb{R}^2$ 不能表示为可列条直线的并集; 更一般地, 空间 $\mathbb{R}^n$ 不能表示为可列个超平面的并集 (注意, 这一结论不能由关于基数的讨论推得, 此因 $\operatorname{card} \mathbb{R}^n = \operatorname{card} \mathbb{R}$; 见定理 1.5-3).

定理 5.3-1(a) 更引人注目的应用是下述结果.

定理 5.1-4 无穷维的 Banach 空间不可能具有可列无穷 Hamel 基 (2.1 节).

特别地, 由单变量或多变量的多项式组成的空间不可能装备范数成为 Banach 空间.

证明 在赋范向量空间 $(X, \|\cdot\|)$ 中给定一组可列无穷 Hamel 基 $(e_j)_{j=0}^{\infty}$, 又由下式定义 X 的子集 F_n:

$$F_n := \operatorname{Span}(e_j)_{j=0}^{n}, \quad \forall n \geqslant 0.$$

首先注意, $X = \bigcup_{n=0}^{\infty} F_n$ 且每一个子集 F_n $(n \geqslant 0)$ 在 X 中都是闭的 (作为 X 的有限维子空间是闭的, 见定理 2.7-1(d)).

接下来我们证明, 对所有 $n \geqslant 0$, $\operatorname{int} F_n = \varnothing$. 实际上, 如果不然, 对某 $n \geqslant 0$, $\operatorname{int} F_n \neq \varnothing$, 则一定存在 $x = \sum_{j=0}^{n} x_j e_j \in F_n$ 及 $r > 0$ 使得 $\overline{B(x;r)} \subset F_n$. 这样, 点

$$y := \frac{r}{\|e_{n+1}\|} e_{n+1} + x$$

就属于 $\overline{B(x;r)}$. 但由于 $(e_j)_{j=0}^{n+1}$ 是线性无关的, y 不可能属于 $F_n = \operatorname{Span}(e_j)_{j=0}^{n}$.

由定理 5.1-3(a) 立得, 空间 X 不可能是完备的.

我们可直接将这一结论应用到 n 个变量的多项式空间 $\mathcal{P}$ 上. 实际上多项式

$$x = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \to x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}, \quad k_i \in \mathbb{N}, \quad 1 \leqslant i \leqslant n$$

就形成空间 $\mathcal{P}$ 的一个可列无穷 Hamel 基. □

习题

5.1-1 这个习题的目的是确立 Cantor 交集定理 (定理 5.1-1) 之逆. 设 (X, d) 是具有以下性质的距离空间: 给定 X 的任一非空闭子集序列 $(A_n)_{n=0}^{\infty}$, 它满足 $A_n \supset A_{n+1}$ 对所有 $n \geqslant 0$ 成立且 $\lim_{n\to\infty} \text{diam } A_n = 0$, 则其交集 $\bigcap_{n=0}^{\infty} A_n$ 定为非空的 (因为 $\lim_{n\to\infty} \text{diam } A_n = 0$, 这个交集由单个元素组成).

证明 (X, d) 是完备的.

5.1-2 设 $f \in \mathcal{C}^{\infty}(\mathbb{R})$ 是具有下述性质的函数: 在每个 $x \in \mathbb{R}$, 存在一个整数 $n(x) \geqslant 0$ 使得 $f^{(n(x))} = 0$. 证明 f 是多项式[2].

5.2 Baire 定理的应用: 连续而无处可微函数的存在性

1872 年 Karl Weierstrass[3] 发表了一个让人觉得不可思议的例子, 它构造出一个 $[0, 1]$ 上的连续函数, 但在 $[0, 1]$ 上处处不可微 (习题 5.2-1). 值得关注的是, 无需给出函数的显式表达, 由 Baire 定理可直接推得这种函数的存在性. 这种处理方式是下一定理讨论的目标, 由定理还可看到, 这种函数的存在有一定的普遍性而不仅仅是一种罕见的例外.

定理 5.2-1 *存在 $[0, 1]$ 上的连续函数, 在 $[0, 1]$ 上处处不可微.*

证明 (i) 设 $f \in \mathcal{C}[0, 1]$ 至少在一点 $a \in [0, 1]$ 处可微. 则有

$$f \in \bigcup_{n=1}^{\infty} F_n,$$

其中

---

[2] 这个令人瞩目的结果属于:

E. COROMINAS; F.S. BALAGUER [1954]: Condiciones para que una función infinitamente derivable sea un polinomio. Revista Matemática Hispano-Americana **14**, 26 – 43.

随后这个结果被扩展到多元函数, 见:

A.B. BOGHOSSIAN; P.D. JOHNSON, JR. [1990]: A pointwise condition for an infinitely differentiable function of several variables to be a polynomial. Journal of Mathematical Analysis and Applications **151**, 17 – 19.

[3] K. WEIERSTRASS [1872]: Über continuirliche Functionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten besitzen. Königliche Akademie der Wissenschaften.

$$F_n := \left\{ f \in \mathcal{C}[0,1];\ \text{存在}\ a \in [0,1]\ \text{使得}\ \sup_{h \neq 0} \left| \frac{f(a+h)-f(a)}{h} \right| \leqslant n \right\}.$$

给定这样一个函数 f, 存在 $h_0 > 0$ 使得

$$\left| \frac{f(a+h)-f(a)}{h} \right| \leqslant \left| \frac{f(a+h)-f(a)}{h} - f'(a) \right| + |f'(a)|$$
$$\leqslant 1 + |f'(a)|, \quad \forall\, 0 < |h| \leqslant h_0.$$

另一方面

$$\left| \frac{f(a+h)-f(a)}{h} \right| \leqslant \frac{2}{h_0} \sup_{0 \leqslant x \leqslant 1} |f(x)|, \quad \forall\, |h| \geqslant h_0.$$

这就得到

$$\sup_{h \neq 0} \left| \frac{f(a+h)-f(a)}{h} \right| < \infty$$

(对于以上及以下的这些不等式, 它们当然是被限制在那些使 $a+h$ 属于区间 $[0,1]$ 的点上). 所以存在 $n_0 \geqslant 1$ 使得

$$\sup_{h \neq 0} \left| \frac{f(a+h)-f(a)}{h} \right| \leqslant n_0.$$

故 $f \in F_{n_0} \subset \bigcup_{n=1}^{\infty} F_n$.

(ii) 每一个集合 F_n $(n \geqslant 1)$ 在 $\mathcal{C}[0,1]$ 中都是闭的. 设函数 $f_k \in F_n$, $k \geqslant 0$, 而 $f \in \mathcal{C}[0,1]$ 使得 $\lim_{k\to\infty} \|f_k - f\| = 0$, 其中 $\|\cdot\|$ 表示空间 $\mathcal{C}[0,1]$ 中的 sup 范数 (在此, 整数 $n \geqslant 1$ 是固定的).

对于每个整数 $k \geqslant 0$, 存在一点 $a_k \in [0,1]$ 使得

$$\sup_{h \neq 0} \left| \frac{f_k(a_k+h)-f_k(a_k)}{h} \right| \leqslant n,$$

此因 $f_k \in F_n$. 由 Bolzano-Weierstrass 性质 (定理 1.4-1(b)), 存在序列 $(a_k)_{k=0}^{\infty}$ 的一个子列 $(a_\ell)_{\ell=0}^{\infty}$ 在区间 $[0,1]$ 中收敛. 记 $a := \lim_{\ell\to\infty} a_\ell$.

给定任意的 $h \neq 0$, 设 h_ℓ 由 $a_\ell + h_\ell = a + h$ 确定. 这样, 就存在一个整数 $\ell_0 = \ell_0(h) \geqslant 0$ 使得 $h_\ell \neq 0$ 对所有 $\ell \geqslant \ell_0$ 成立. 利用以下三个关系式

$$|f_\ell(a_\ell+h_\ell) - f(a+h)| = |f_\ell(a_\ell+h_\ell) - f(a_\ell+h_\ell)| \leqslant \|f_\ell - f\|,$$
$$|f_\ell(a_\ell) - f(a)| \leqslant |f_\ell(a_\ell) - f(a_\ell)| + |f(a_\ell) - f(a)|$$
$$\leqslant \|f_\ell - f\| + |f(a_\ell) - f(a)|,$$
$$\lim_{\ell\to\infty} h_\ell = h,$$

并考虑到函数 f 的连续性, 得到

$$\left|\frac{f(a+h)-f(a)}{h}\right| = \lim_{\substack{\ell\to\infty \\ \ell\geqslant\ell_0}} \left|\frac{f_\ell(a_\ell+h_\ell)-f_\ell(a_\ell)}{h_\ell}\right| \leqslant n,$$

此因对所有 $\ell\geqslant\ell_0$, 有 $f_\ell\in F_n$. 这就证明了 $f\in F_n$, 所以 F_n 是闭的.

(iii) 每一个集合 F_n $(n\geqslant 1)$ 的内部均为空集. 这就是要证明, 给定任意的函数 $f\in F_n$ 及任意的 $\varepsilon>0$, 存在一个函数 $g\in\mathcal{C}[0,1]$ 使得 $\|g-f\|\leqslant\varepsilon$ 且 $g\notin F_n$ (整数 $n\geqslant 1$ 在此也是固定的).

为此, 我们首先注意, 由 Weierstrass 逼近定理 (定理 2.13-3), 存在一个多项式 $p=p(f,\varepsilon)$ 使得 $\|f-p\|\leqslant\varepsilon/2$. 给定这样一个多项式 p 后, 我们 (譬如说, 从 $(0,p(0))$ 开始) 构造一个分段仿射函数 $g=g(p)=g(f,\varepsilon)\in\mathcal{C}[0,1]$, 使得 $\|g-p\|\leqslant\varepsilon/2$ 且在其导数 $g'(x)$ 有定义的所有点 $x\in[0,1]$ 上成立 $|g'(x)|>n$ (见图 5.2-1). 要注意, 这样一个函数 g 的存在关键在于 $\sup_{0\leqslant x\leqslant 1}|p'(x)|<\infty$ 这一事实, 此因 p 是多项式.

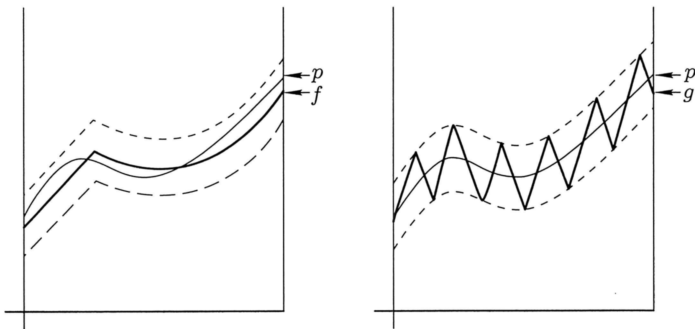


图 5.2-1　定理 5.2-1 的证明中分段仿射函数 g 的构造

如此构造的函数 g 显然具有所要求的性质.

(iv) Baire 定理 (定理 5.1-2) 意味着

$$\operatorname{int}\left(\bigcup_{n=1}^{\infty} F_n\right)=\varnothing,$$

此因空间 $\mathcal{C}[0,1]$ 是完备的 (定理 3.2-2). 这就得到 $\mathcal{C}[0,1]-(\bigcup_{n=1}^{\infty}F_n)\neq\varnothing$, 即存在函数, 它们在 $[0,1]$ 上连续但在 $[0,1]$ 上处处不可微. □

实际上, 在上述证明结尾所建立的关系式 $\operatorname{int}(\bigcup_{n=1}^{\infty}F_n)=\varnothing$ 包含着更多的内容, 即给定任意一个至少在一点可微的函数 $f\in\mathcal{C}[0,1]$, 从而 $f\in\bigcup_{n=1}^{\infty}F_n$, 并给定任意 $\varepsilon>0$, 则存在一个在 $[0,1]$ 上处处不可微的函数 $g\in\mathcal{C}[0,1]$ 使得 $\|f-g\|\leqslant\varepsilon$. 这说明, 任何一个在 $[0,1]$ 上连续、至少在 $[0,1]$ 中一点可微的函数, 是一个无处可微连续函数序列的一致极限. 换言之, 实际上有 "很多" 连续函数是处处不可微的.

习题

5.2-1 证明由

$$f : x \in \mathbb{R} \to f(x) := \sum_{n=0}^{\infty} \frac{1}{2^n} \sin(3^n x)$$

给出的 Weierstrass 函数 $f : \mathbb{R} \to \mathbb{R}$ 在 $\mathbb{R}$ 上是有定义且连续的, 但处处不可微.

5.2-2 证明由

$$f : x \in \mathbb{R} \to f(x) := \sum_{n=0}^{\infty} \frac{1}{n^2} \sin(n^2 \pi x)$$

定义的 Hardy 函数[4] $f : \mathbb{R} \to \mathbb{R}$ 在 $\mathbb{R}$ 上是有定义且连续的, 但处处不可微.

5.3 Banach-Steinhaus 定理, 即一致有界性原理; 对数值求积公式的应用

给定两个赋范向量空间 X 与 Y, 考察按下述意义是 "一致有界" 的连续线性算子 $A_i \in \mathcal{L}(X;Y)$ 构成的算子族 $(A_i)_{i\in I}$:

$$\sup_{i\in I} \|A_i\|_{\mathcal{L}(X,Y)} < \infty.$$

因为 $\|A_i x\| \leqslant \|A_i\|\|x\|$ 对所有 $i \in I$ 及所有 $x \in X$ 成立, 所以必定有

$$\text{对每一个 } x \in X, \quad \sup_{i\in I} \|A_i x\|_Y < \infty.$$

值得关注的是, 如果空间 X 是完备的 (这个假定是实质性的, 见习题 5.3-1), 这个对映射 A_i, $i \in I$ (在上述意义下的) 一致有界的必要条件就变成充分的了. 这个 "一致有界性原理" 是 Baire 定理的一个推论.

定理 5.3-1 (Banach-Steinhaus 定理[5], 即一致有界性定理) 设 X 是 Banach 空间, Y 是赋范向量空间, 而 $(A_i)_{i\in I}$ 为映射 $A_i \in \mathcal{L}(X;Y)$ 构成的算子族, 满足

$$\text{对每一个 } x \in X, \quad \sup_{x\in I} \|A_i x\|_Y < \infty.$$

则

$$\sup_{i\in I} \|A_i\|_{\mathcal{L}(X;Y)} < \infty.$$

---

[4] G.H. Hardy [1916]: Weierstrass' Non-differentable Function. Transactions, American Mathematical Society **17**, 301 – 325.

[5] S. Banach; H. Steinhaus [1927]: Sur le principe de la condensation de singularités. Fundamenta Mathematicae **9**, 50 – 61.

证明 同一个符号 $\|\cdot\|$ 在本证明中可能表示所用到的不同范数. 对每个整数 $n \geqslant 0$, 定义集合

$$F_n := \{x \in X; \sup_{i\in I} \|A_i x\| \leqslant n\}.$$

给定任意 $x \in X$, 由假设 $\sup_{i\in I} \|A_i x\| < \infty$, 故存在一个整数 $n(x) \geqslant 0$ 使得 $\sup_{i\in I} \|A_i x\| \leqslant n(x)$. 这意味着 $x \in F_{n(x)}$. 所以

$$X = \bigcup_{n=0}^{\infty} F_n.$$

根据定义, $x \in F_n$ 的充分必要条件是 $\|A_i x\| \leqslant n$ 对所有 $i \in I$ 成立, 或等价地, $x \in \{z \in X; \|A_i z\| \leqslant n\}$ 对所有 $i \in I$ 成立. 所以, 集合 F_n 又可以表示为

$$F_n = \bigcap_{i\in I} \{z \in X; \|A_i z\| \leqslant n\},$$

故作为 X 中闭子集 (由假定, 每一线性算子 A_i 都是连续的) 的交集, F_n 是 X 中的闭集.

因为 X 是完备的, Baire 定理适用, 故存在一个整数 $n_0 \geqslant 0$ 使得 $\operatorname{int} F_{n_0} \neq \varnothing$ (定理 5.1-3(b)). 所以存在 $x_0 \in F_{n_0}$ 及 $r > 0$ 使得 $\overline{B(x_0; r)} \subset F_{n_0}$. 由 F_{n_0} 的定义, 这意味着

$$\|A_i z\| \leqslant n_0, \quad \text{对所有 } z \in \overline{B(x_0; r)} \text{ 及 } i \in I.$$

由于任一个非零向量 $x \in X$ 均可以写成

$$x = \frac{\|x\|}{r}(z - x_0), \quad \text{其中 } z := \left(x_0 + \frac{r}{\|x\|}x\right) \in \overline{B(x_0; r)},$$

这就有

$$\begin{aligned} \|A_i x\| &\leqslant \frac{\|x\|}{r}(\|A_i z\| + \|A_i x_0\|) \\ &\leqslant \frac{1}{r}(n_0 + \|A_i x_0\|)\|x\| \\ &\leqslant \frac{1}{r}\left(n_0 + \sup_{i\in I}\|A_i x_0\|\right)\|x\|, \quad \text{对所有 } i \in I \text{ 及 } x \in X. \end{aligned}$$

由假设, $\sup_{i\in I} \|A_i x_0\| < \infty$, 所以我们有

$$\sup_{i\in I} \|A_i\| \leqslant \frac{1}{r}\left(n_0 + \sup_{i\in I}\|A_i x_0\|\right) < \infty. \qquad \square$$

Banach-Steinhaus 定理常被用以下述推论的形式. 下面我们将称其为 Banach-Steinhaus 定理的推论.

定理 5.3-2 (Banach-Steinhaus 定理的推论) 设 X 是 Banach 空间, Y 是赋范向量空间, 而 $(A_n)_{n=1}^{\infty}$ 为映射 $A_n \in \mathcal{L}(X;Y)$ 构成的算子族, 使得对每一 $x \in X$, 序列 $(A_n x)_{n=1}^{\infty}$ 在 Y 中收敛. 则

$$\sup_{n \geqslant 1} \|A_n\| < \infty.$$

进而, 设映射 $A: X \to Y$ 由下式定义:

$$Ax := \lim_{n\to\infty} A_n x, \quad 对每个\ x \in X.$$

则

$$A \in \mathcal{L}(X;Y) \quad 且 \quad \|A\| \leqslant \liminf_{n\to\infty} \|A_n\|.$$

证明　每一个序列 $(A_n x)_{n=1}^{\infty}$ 的收敛性意味着

$$对每个\ x \in X, \quad \sup_{n \geqslant 1} \|A_n x\| < \infty.$$

由 Banach-Steinhaus 定理 (定理 5.3-1) 即有

$$\sup_{n \geqslant 1} \|A_n\| < \infty.$$

由于每个映射 A_n 都是线性的, 再注意到加法与数乘的连续性 (见定理 2.2-5) 就可以得到, 对每个 $x \in X$, 由 $Ax := \lim_{n\to\infty} A_n x$ 定义的映射 $A: X \to Y$ 是线性的. 此外, 给定非零向量 $x \in X$,

$$\begin{aligned}\frac{\|Ax\|}{\|x\|} &= \lim_{n\to\infty} \frac{\|A_n x\|}{\|x\|} = \liminf_{n\to\infty} \frac{\|A_n x\|}{\|x\|} \\ &\leqslant \liminf_{n\to\infty} \|A_n\| < \infty,\end{aligned}$$

此因对所有 $n \geqslant 1$, 均有 $\dfrac{\|A_n x\|}{\|x\|} \leqslant \|A_n\|$. 这就得到 $A \in \mathcal{L}(X;Y)$ 而且

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \leqslant \liminf_{n\to\infty} \|A_n\|. \qquad \square$$

注　(1) 有时也称定理 5.3-2 为 Banach-Steinhaus 定理.

(2) 仅在定理 5.3-2 的假设下, 关于序列 $(A_n)_{n=1}^{\infty}$ 在空间 $\mathcal{L}(X;Y)$ 中是否收敛于 A, 得不到任何结论.

作为 Banach-Steinhaus 定理威力的第一个例子, 我们说明如何利用这一定理给出一个判定一大类数值求积公式收敛性的绝妙准则 (定理 5.3-3).

更明确地说, 给定一个权函数 $\omega \in L^1(0,1)$, 我们的目标是, 对任意函数 $f \in \mathcal{C}[0,1]$, 用 "容易计算" 的有限和 "尽可能精确地" 逼近积分

$$\ell(f) := \int_0^1 f(x)\omega(x)dx$$

(在此区间选定为 $[0,1]$ 当然只是为了确定起见). 达此目标的一个自然的方式是, 对每个 $n \geqslant 0$, 适当选取 $n+1$ 个不同的节点 $0 \leqslant x_0^n < x_1^n < \cdots < x_n^n \leqslant 1$ 及 $n+1$ 个权 $\omega_j^n \in \mathbb{R}$, $0 \leqslant j \leqslant n$, 然后用数值求积公式

$$\ell_n(f) := \sum_{j=0}^{n} \omega_j^n f(x_j^n)$$

来逼近积分 $\ell(f)$. 注意, 在空间 $\mathcal{C}[0,1]$ 装备 sup 范数后, 上面定义的映射 $\ell : \mathcal{C}[0,1] \to \mathbb{R}$ 及 $\ell_n : \mathcal{C}[0,1] \to \mathbb{R}$ 显然是连续线性泛函.

有各种各样的构造数值求积公式的方法[6]. 例如, 设 $\sum_{j=0}^{n} f(x_j^n) p_j^n \in \mathcal{P}_n[0,1]$ 表示函数 $f \in \mathcal{C}[0,1]$ 的阶数 $\leqslant n$ 的 Lagrange 插值多项式, 插值中采用的是等距节点 $x_j^n = j/n$, $0 \leqslant j \leqslant n$ (在 5.4 节中将给出这种多项式的定义). 一种逼近积分 $\ell(f)$ 的方法是利用 Newton-Cotes 求积公式:

$$\ell_n(f) := \sum_{j=0}^{n} \left(\int_0^1 p_j^n(x) \omega(x) dx \right) f(x_j^n).$$

由其构造可以看出, 对所有阶数 $\leqslant n$ 的多项式而言, 这个公式是严格精确的.

一个 (可能更有效的) 方式是, 考虑同时选取 $n+1$ 个节点 x_j^n 及 $n+1$ 个权值 ω_j^n $(0 \leqslant j \leqslant n)$, 使得所给出的数值求积公式对所有阶数 $\leqslant 2n+1$ 的多项式是严格精确的. 这是否能够做到并不是显然的, 因为这要求解关于未知量 x_j^n 及 ω_j^n $(0 \leqslant j \leqslant n)$ 的 $2n+2$ 个非线性方程构成的方程组. 然而, 由于这个问题与关于严格正权函数正交的多项式 (见习题 5.3-3) 的性质之间存在一种特殊的联系, 使得求解这一非线性方程组成为可能. 而相应的公式 $\ell_n(f)$ 称为 Gauss-Jacobi 求积公式.

下面的定理显示, 对于任意的 $f \in \mathcal{C}[0,1]$, 可以给出一个当 $n \to \infty$ 时, 一般形式数值求积公式 $\ell_n(f)$ 序列收敛于积分 $\ell(f)$ 的极其简明的充分必要条件. 注意, 对于节点 x_j^n, $0 \leqslant j \leqslant n$, 除了要求它们各异之外, 未做任何其他假设.

定理 5.3-3 (Polya 定理[7])　给定一个权函数 $\omega \in L^1(0,1)$, 及一个连续线性泛函 $\ell_n : \mathcal{C}[0,1] \to \mathbb{R}$ $(n \geqslant 0)$ 的形如

$$\ell_n : f \in \mathcal{C}[0,1] \to \ell_n(f) := \sum_{j=0}^{n} \omega_j^n f(x_j^n) \in \mathbb{R},$$

$$\text{其中 } 0 \leqslant x_0^n < x_1^n < \cdots < x_n^n \leqslant 1$$

的序列. 设该序列具有以下性质:

$$\lim_{n\to\infty} \left| \int_0^1 p(x)\omega(x)dx - \ell_n(p) \right| = 0, \quad \text{对任意 } p \in \mathcal{P}[0,1].$$

[6] 关于数值求积深入的讨论可在 Davis & Rabinowitz [1975] 中找到.

[7] G. Pólya [1933]: Über die Konvergenz von Quadraturverfahren. Mathematische Zeitschrift **37**, 264 – 286.

则

$$\lim_{n\to\infty}\left|\int_0^1 f(x)\omega(x)dx-\ell_n(f)\right|=0,\quad 对任意\ f\in\mathcal{C}[0,1]$$

的充分必要条件是

$$\sup_{n\geqslant 0}\left(\sum_{j=0}^{n}|\omega_j^n|\right)<\infty.$$

证明 显然, 对所有 $f\in\mathcal{C}[0,1]$ 有 $|\ell_n(f)|\leqslant\left(\sum_{j=0}^{n}|\omega_j^n|\right)\|f\|$, 故

$$\|\ell_n\|=\sup_{f\neq 0}\frac{|\ell_n(f)|}{\|f\|}\leqslant\sum_{j=0}^{n}|\omega_j^n|.$$

设 $f_0\in\mathcal{C}[0,1]$ 表示如下定义的分段仿射连续函数

$$f_0(0)=\operatorname{sgn}\omega_0^n,\quad f_0(x_j^n)=\operatorname{sgn}\omega_j^n,\quad 0\leqslant j\leqslant n,\quad 而\ f_0(1)=\operatorname{sgn}\omega_n^n.$$

则有 $|\ell_n(f_0)|=\sum_{j=0}^{n}|\omega_j^n|$ 及 $\|f_0\|=1$. 这样就得

$$\|\ell_n\|\geqslant\frac{|\ell_n(f_0)|}{\|f_0\|}=\sum_{j=0}^{n}|\omega_j^n|,\quad 对每个\ n\geqslant 0.$$

所以 $\|\ell_n\|=\sum_{j=0}^{n}|\omega_j^n|$.

由 Banach-Steinhaus 定理的推论 (定理 5.3-2) 得 $\sup_{n\geqslant 0}\|\ell_n\|<\infty$, 所以必要性部分得证.

反之, 假定 $\sup_{n\geqslant 0}\left(\sum_{j=0}^{\infty}|\omega_j^n|\right)=\sup_{n\geqslant 0}\|\ell_n\|<\infty$. 对任意的 $f\in\mathcal{C}[0,1]$ 及任意的 $p\in\mathcal{P}[0,1]$, 我们有

$$\begin{aligned}&\left|\ell_n(f)-\int_0^1 f(x)\omega(x)dx\right|\\ &\leqslant|\ell_n(f-p)|+\left|\ell_n(p)-\int_0^1 p(x)\omega(x)dx\right|+\left|\int_0^1(f(x)-p(x))\omega(x)dx\right|\\ &\leqslant\left(\sup_{n\geqslant 0}\|\ell_n\|+\|\omega\|_{L^1(0,1)}\right)\|f-p\|+\left|\ell_n(p)-\int_0^1 p(x)\omega(x)dx\right|.\end{aligned}$$

给定任意的 $f\in\mathcal{C}[0,1]$ 及任意的 $\varepsilon>0$, Weierstrass 逼近定理 (定理 2.12-3) 说明, 存在一个多项式 $p=p(f;\varepsilon)\in\mathcal{P}[0,1]$ 使得

$$\left(\sup_{n\geqslant 0}\|\ell_n\|+\|\omega\|_{L^1(0,1)}\right)\|f-p\|\leqslant\frac{\varepsilon}{2}.$$

由定理假设知, 存在 $n_0 = n_0(p) = n_0(f;\varepsilon)$ 使得

$$\left|\ell_n(p) - \int_0^1 p(x)\omega(x)dx\right| \leqslant \frac{\varepsilon}{2}, \quad 对所有 n \geqslant n_0,$$

所以有

$$\left|\ell_n(f) - \int_0^1 f(x)\omega(x)dx\right| \leqslant \varepsilon, \quad 对所有 n \geqslant n_0.$$

这就证明了充分性部分. □

注 (1) 与必要性部分的证明不同, 在充分性部分的证明中没用到 Banach-Steinhaus 定理.

(2) Newton-Cotes 及 Gauss-Jacobi 求积公式显然都满足性质

$$\lim_{n\to\infty}\left|\ell_n(p) - \int_0^1 p(x)\omega(x)dx\right| = 0, \quad 对任意多项式 p \in \mathcal{P}[0,1],$$

此因给定任意的多项式 $p \in \mathcal{P}[0,1]$, 存在一个整数 $n=n(p)$ 使得 $\ell_n(p)=\int_0^1 p(x)\omega(x)dx$.

习题

5.3-1 这个习题给出了当空间不完备时关于 Banach-Steinhaus 定理的一个反例. 下面, 符号 $\mathcal{P}$ 表示所有实多项式的空间.

(1) 给定一个多项式 $p : x \in \mathbb{R} \to \sum_{k=0}^{m} c_k x^k$, 设 $\|p\| := \max_{0\leqslant k\leqslant m} |c_k|$. 证明 $\|\cdot\|$ 定义空间 $\mathcal{P}$ 的一个范数.

(2) 直接证明 $(\mathcal{P}, \|\cdot\|)$ 是不完备的 (即不利用定理 5.1-4).

(3) 对每个 $n \geqslant 0$, 以 $A_n p := \sum_{k=0}^{\min\{m,n\}} c_k$ 定义线性算子 $A_n : \mathcal{P} \to \mathbb{R}$. 证明每个算子 A_n 都是连续的.

(4) 证明 $\sup_{n\geqslant 0} |A_n p| < \infty$ 对每个 $p \in \mathcal{P}$ 都成立, 但 $\sup_{n\geqslant 0} \|A_n\| = \infty$.

5.3-2 (1) 设 X 是 Banach 空间, Y 及 Z 是赋范向量空间, 而 $B : X \times Y \to Z$ 是一个双线性映射, 它在下述意义下是“分立”连续的:

$$\begin{aligned}&对每个\ y \in Y,\ \lim_{n\to\infty} x_n = x\ 在\ X\ 中成立意味着\\ &\qquad \lim_{n\to\infty} B(x_n, y) = B(x, y)\ 在\ Z\ 中成立;\\ &对每个\ x \in X,\ \lim_{n\to\infty} y_n = y\ 在\ Y\ 中成立意味着\\ &\qquad \lim_{n\to\infty} B(x, y_n) = B(x, y)\ 在\ Z\ 中成立.\end{aligned}$$

利用 Banach-Steinhaus 定理证明 B 是连续的, 即对每个 $(x, y) \in X \times Y$,

$$\begin{aligned}&\lim_{n\to\infty} x_n = x\ 在\ X\ 中成立且\ \lim_{n\to\infty} y_n = y\ 在\ Y\ 中成立意味着\\ &\qquad \lim_{n\to\infty} B(x_n, y_n) = B(x, y)\ 在\ Z\ 中成立.\end{aligned}$$

(2) 给出一个例子, 其中 X, Y, Z 是赋范向量空间, $B: X \times Y \to Z$ 为分立连续的双线性映射, 但它不连续.

5.3-3 给定一个函数 $\omega \in L^1(0,1)$ 且在 $[0,1]$ 中几乎处处满足条件 $\omega > 0$. 设 p_n $(n \geqslant 0)$ 表示关于权函数 ω 正交的多项式 (见习题 4.8-3).

(1) 对每个整数 $n \geqslant 1$, 令 x_j^n $(0 \leqslant j \leqslant n)$ 表示 p_n 的零点 (这些零点都是单重实根, 而且全在开区间 $]0,1[$ 中; 见前注). 试证明, 存在常数 ω_j^n, $0 \leqslant j \leqslant n$, 满足

$$\omega_j^n > 0, \quad 0 \leqslant j \leqslant n, \quad 且 \sum_{j=0}^{n} \omega_j^n p(x_j^n) = 0, \quad 对所有 p \in \mathcal{P}_{2n+1}[0,1].$$

(2) 证明, 如果 $f \in \mathcal{C}^{2n+2}[0,1]$, 则存在一点 $\xi = \xi(f) \in]0,1[$ 使得

$$\int_0^1 f(x)\omega(x)dx - \sum_{j=0}^{n} \omega_j^n f(x_j^n) = \frac{1}{(2n+2)!} f^{(2n+2)}(\xi).$$

因此如果 $f \in \mathcal{C}^\infty[0,1]$ 而且 $\lim_{n\to\infty}\left(\dfrac{1}{n!}\sup_{0\leqslant x\leqslant 1}|f^{(n)}(x)|\right) = 0$, 则有

$$\lim_{n\to\infty}\left(\sum_{j=0}^{n} \omega_j^n f(x_j^n)\right) = \int_0^1 f(x)\omega(x)dx.$$

5.3-4 给定一个权函数 $\omega \in L^1(0,1)$ 及一个线性连续泛函 $\ell_n : \mathcal{C}[0,1] \to \mathbb{R}$ $(n \geqslant 0)$ 的形如

$$\ell_n : f \in \mathcal{C}[0,1] \to \ell_n(f) := \sum_{j=0}^{n} \omega_j^n f(x_j^n) \in \mathbb{R},$$
$$其中\ 0 \leqslant x_0^n < x_1^n < \cdots < x_n^n \leqslant 1$$

的序列. 该序列具有以下性质:

$$\lim_{n\to\infty}\left|\int_0^1 p(x)\omega(x)dx - \ell_n(p)\right| = 0, \quad 对任一\ p \in \mathcal{P}[0,1].$$

证明, 如果 $\omega_j^n \geqslant 0$ 对所有 $n \geqslant 0$ 及所有 $0 \leqslant j \leqslant n$ 成立, 则

$$\lim_{n\to\infty}\left|\int_0^1 f(x)\omega(x)dx - \ell_n(f)\right| = 0, \quad 对任一\ f \in \mathcal{C}[0,1].$$

这个结果称为 Steklov 定理[8].

5.4 Banach-Steinhaus 定理的应用: Lagrange 插值的发散性

下面, $[a,b]$ 是一紧区间, $[a,b] \subset \mathbb{R}$, $a < b$, 而 $\mathcal{C}[a,b]$ 表示所有连续函数 $f : [a,b] \to \mathbb{R}$ 形成的 Banach 空间, 其中装备以由 $\|f\| = \sup_{a\leqslant x\leqslant b}|f(x)|$ 定义的 sup 范数. 对

[8] 冠名源自 Vladimir Andreevich Steklov (1864—1926).

于每个 $n \geqslant 0$, $\mathcal{P}_n$ 表示所有阶数 $\leqslant n$ 的单实变量多项式空间, 而 $\mathcal{P}_n[a,b]$ 表示所有 $p \in \mathcal{P}_n$ 在 $[a,b]$ 上的限制形成的 $\mathcal{C}[a,b]$ 的子空间.

下一定理描述熟知的 Lagrange 插值[9], 就是利用多项式在有限个点上内插一个给定的函数. 更一般的 Hermite 插值[10] 则是另外还要在有限个点上内插给定函数的一些导数 (Hermite 插值的例子在习题 5.4-2 及 5.4-3 中给出). 多变量的 Lagrange 插值将在 7.11 节中予以讨论.

定理 5.4-1 对每个整数 $n \geqslant 0$, 给定任意的 $n+1$ 个不同的点 $a \leqslant x_0 < x_1 < \cdots < x_n \leqslant b$. 则给定任意函数 $f \in \mathcal{C}[a,b]$, 存在一个且仅有一个多项式 $L_n f \in \mathcal{P}_n[a,b]$ 满足

$$L_n f(x_i) = f(x_i), \quad 0 \leqslant i \leqslant n.$$

这个多项式称为 f 的关于 $n+1$ 个节点 $x_i \in [a,b]$, 阶数 $\leqslant n$ 的 Lagrange 插值多项式. 它由下式给出:

$$L_n f(x) = \sum_{j=0}^{n} f(x_j) p_j(x), \quad a \leqslant x \leqslant b,$$

其中 $n+1$ 个多项式 $p_j \in \mathcal{P}_n[a,b]$, $0 \leqslant j \leqslant n$, 由

$$p_j(x) := \prod_{\begin{cases} i = 0 \\ i \neq j \end{cases}}^{n} \frac{x - x_i}{x_j - x_i}, \quad a \leqslant x \leqslant b,$$

确定.

由这种方式定义的算子 $L_n : \mathcal{C}[a,b] \to \mathcal{C}[a,b]$ 是线性的, 关于范数

$$\|L_n\| = \sup_{a \leqslant x \leqslant b} \left(\sum_{j=0}^{n} |p_j(x)| \right)$$

是连续的, 而且满足

$$L_n p = p, \quad \text{对所有 } p \in \mathcal{P}_n[a,b].$$

证明 关系式 $p_j(x_i) = \delta_{ij}$, $0 \leqslant i, j \leqslant n$, 说明给定任意函数 $f \in \mathcal{C}[a,b]$, 由其确定的多项式 $p := \sum_{j=0}^{n} f(x_j) p_j \in \mathcal{P}_n[a,b]$ 满足 $p(x_i) = f(x_i)$, $0 \leqslant i \leqslant n$.

要求得这样的插值多项式, 实际上就是求解一个有 $n+1$ 个 (即节点数) 方程, 含有同样数量的未知量 (即未知多项式关于 $\mathcal{P}_n[a,b]$ 的标准基的系数) 的线性方程组. 但由系数阵为方阵的线性方程组熟知的性质可断定, 解的存在性 (前面已证明) 意味着唯

---

[9] 冠名源自 J.-L. LAGRANGE [1812]: Leçons élémentaires de mathématiques données à l'Ecole Normale en 1795. Journal de l'Ecole Polytechnique, VIIe et VIIIe cahiers, t-II.

[10] 冠名源自 C. HERMITE [1878]: Sur la formule d'interpolation de Lagrange. Journal für die reine und angewandte Mathematik **84**, 70 – 79.

一性. 因此 f 的阶数 $\leqslant n$ 的唯一插值多项式由下式给出:

$$L_n f = \sum_{j=0}^{n} f(x_j) p_j.$$

这也证明了, $n+1$ 个多项式 p_j $(0 \leqslant j \leqslant n)$ 形成 $\mathcal{P}_n$ 的一组基.

显然, 以这种方式定义的算子 $L_n : \mathcal{C}[a,b] \to \mathcal{C}[a,b]$ 是线性的, 而且 $L_n p = p$ 对所有 $p \in \mathcal{P}_n[a,b]$ 成立 (插值多项式是唯一的). 由表达式 $L_n f = \sum_{j=0}^{n} f(x_j) p_j$, 立即可以得到

$$\|L_n\| \leqslant \sup_{a \leqslant x \leqslant b} \left(\sum_{j=0}^{n} |p_j(x)| \right).$$

设 $\zeta \in [a,b]$ 使

$$\sum_{j=0}^{n} |p_j(\zeta)| = \sup_{a \leqslant x \leqslant b} \left(\sum_{j=0}^{n} |p_j(x)| \right),$$

又设 $\widetilde{f} \in \mathcal{C}[a,b]$ 表示由下式确定的分段仿射连续函数:

$$\widetilde{f}(a) = \operatorname{sgn} p_0(\zeta), \quad \widetilde{f}(x_j) = \operatorname{sgn} p_j(\zeta), \quad 0 \leqslant j \leqslant n, \quad \widetilde{f}(b) = \operatorname{sgn} p_n(\zeta).$$

设函数 $f_0 \in \mathcal{C}[a,b]$ 定义为 $f_0(x) := 1$, $a \leqslant x \leqslant b$. 则 $L_n f_0 = f_0$, 这意味着

$$\sum_{j=0}^{n} p_j(x) = 1, \quad a \leqslant x \leqslant b.$$

$\widetilde{f}$ 的定义说明, 要么 $\|\widetilde{f}\| = 0$, 要么 $\|\widetilde{f}\| = 1$; 但是因为 $\sum_{j=0}^{n} p_j(\zeta) = 1$ (故这些 $p_j(\zeta)$, $0 \leqslant j \leqslant n$, 不可能同时取零值), $\|\widetilde{f}\| = 0$ 不可能成立, 所以 $\|\widetilde{f}\| = 1$. 此外

$$\|L_n \widetilde{f}\| \geqslant |L_n \widetilde{f}(\zeta)| = \left| \sum_{j=0}^{n} \widetilde{f}(x_j) p_j(\zeta) \right| = \sum_{j=0}^{n} |p_j(\zeta)|.$$

利用这些关系式, 并注意到

$$\|L_n\| = \sup_{f \neq 0} \frac{\|L_n f\|}{\|f\|} \geqslant \frac{\|L_n \widetilde{f}\|}{\|\widetilde{f}\|},$$

就得到 $\|L_n\| \geqslant \sup_{a \leqslant x \leqslant b} \left(\sum_{j=0}^{n} |p_j(x)| \right)$. 所以

$$\|L_n\| = \sup_{a \leqslant x \leqslant b} \left(\sum_{j=0}^{n} |p_j(x)| \right).$$

□

一个自然的问题随即产生: 在关于函数 $f \in \mathcal{C}[a,b]$ 的何种假设下, Lagrange 插值是一个收敛的逼近格式, 即 $\lim_{n\to\infty} \|L_n f - f\| = 0$? 虽然不难证明, 如果函数 f 是无穷次可微的, 且其导数"不增长得太快", 收敛性成立 (习题 5.4-1), 但 Sergei Natanovich Bernstein 于 1918 年给出著名的反例后, 下述事实就为人所知了: 函数 $f: x \in [-1,1] \to |x|$, 在区间 $[0,1]$ 以 -1, 0 及 1 为特定节点的等距节点上, 除了在点 -1, 0 及 1 外, 并不点点收敛于 f.

注 可以证明[11] $\lim_{n\to\infty} \left|\dfrac{L_n f(x) - |x|}{\prod_{i=0}^n (x - x_i)}\right|^{\frac{1}{2}} = e$ 对所有 $x \in \mathbb{R}$ 成立.

值得关注的是, 借助于 Banach-Steinhaus 定理, 不需要给出这类函数的任何显式表示, 即可断定对于某些连续函数其 Lagrange 插值的发散性. 注意, 除了对每个 $n \geqslant 0$ 要求节点 x_i^n $(0 \leqslant i \leqslant n)$ 是各异的, 对它们不做任何其他假设.

定理 5.4-2 对每个整数 $n \geqslant 0$, 设给定 $n+1$ 个各异的节点 $a \leqslant x_0^n < x_1^n < \cdots < x_n^n \leqslant b$. 给定任意函数 $f \in \mathcal{C}[a,b]$, 设对任意 $n \geqslant 0$, 其 Lagrange 插值多项式 $L_n f \in \mathcal{P}_n[a,b]$ 由 $L_n f(x_i^n) = f(x_i^n)$ $(0 \leqslant i \leqslant n)$ 定义. 则

$$\sup_{n\geqslant 0} \|L_n f\| = \infty, \quad \text{对某些 } f \in \mathcal{C}[a,b],$$

当然这一性质更断言, 对于这种函数, $L_n f$ 不可能一致收敛于 f.

证明 在定理 5.4-1 中得到

$$\|L_n\| = \sup_{a\leqslant x\leqslant b} \left(\sum_{j=0}^n |p_j^n(x)|\right),$$

其中

$$p_j^n(x) := \prod_{\begin{cases} i=0 \\ i\neq j\end{cases}}^n \left(\frac{x - x_i^n}{x_j^n - x_i^n}\right), \quad a \leqslant x \leqslant b,$$

而且可以证明

$$\lim_{n\to\infty} \|L_n\| = \infty$$

(见习题 5.4-4).

如果我们有 $\sup_{n\geqslant 0} \|L_n f\| < \infty$ 对所有 $f \in \mathcal{C}[a,b]$ 成立, 那么由 Banach-Steinhaus 定理 (定理 5.3-1) 得 $\sup_{n\geqslant 0} \|L_n\| < \infty$, 这与 $\lim_{n\to\infty} \|L_n\| = \infty$ 矛盾. 所以至少对某一个函数 $f \in \mathcal{C}[a,b]$, 有 $\sup_{n\geqslant 0} \|L_n f\| = \infty$. □

注 如果代替 Banach-Steinhaus 定理, 我们利用它的推论 (定理 5.3-2), 我们也能得到如下结论: 存在连续函数 f, 其 Lagrange 插值多项式不一致收敛于 f. 但我们得不到 $\sup_{n\geqslant 1} \|L_n f\| = \infty$ 这一结论.

---

[11] X. LI; R.N. MOHAPATRA [1993]: On the convergence of Lagrange interpolation with equidistant nodes. Proceedings of the American Mathematical Society **118**, 1205 – 1212.

出现在上述证明中的范数 $\|L_n\|$ 称为 Lebesgue 常数[12].

作为这一节的结束, 对于紧区间上连续函数的多项式逼近问题, 我们给予某种一般的讨论.

在这方面 "理想的目标" 是, 寻求具有以下四条性质的映射 $A_n : \mathcal{C}[a,b] \to \mathcal{P}_n[a,b]$ 的序列 $(A_n)_{n=0}^{\infty}$: 算子 A_n 是线性且连续的, 算子作用在阶数 $\leqslant n$ 的所有多项式上均保持不变, 即 $A_n p = p$ 对所有 $p \in \mathcal{P}_n[a,b]$ 及 $n \geqslant 1$ 成立 (与 Lagrange 插值关联的算子 L_n 就满足这三条性质), 而且 (最重要的) 它们还满足

$$\lim_{n\to\infty} \|A_n f - f\| = 0, \quad 对所有\ f \in \mathcal{C}[a,b].$$

但是根据下面出色的结果 (如所预料的, 其证明又依赖于 Banach-Steinhaus 定理), 这个目标是不可能达到的. 以下定理说明, 在定理 5.4-2 中确立的发散现象决不只局限于 Lagrange 插值.

定理 5.4-3 (Kharshiladze-Lozinski 逼近定理[13]) 对任何一个由线性连续且作用在阶数 $\leqslant n$ 的多项式上保持不变的映射 $A_n : \mathcal{C}[a,b] \to \mathcal{P}_n[a,b] \subset \mathcal{C}[a,b]$ 组成的序列 $(A_n)_{n=0}^{\infty}$, 一定存在某 $f \in \mathcal{C}[a,b]$ 使得

$$\lim_{n\to\infty} \sup \|A_n f - f\| = \infty,$$

这一性质更说明, 对于这种函数 f, $A_n f$ 不可能一致收敛于 f.

根据这个负面的结果, 重温一下其他类型的多项式逼近是有意义的 (为确定起见, 在区间 $[a,b] = [0,1]$ 上讨论).

首先考虑 Bernstein 多项式 $B_n f \in \mathcal{P}_n[0,1]$, $n \geqslant 0$. 对任意函数 $f \in \mathcal{C}[0,1]$, 它由下式定义 (见定理 2.13-2):

$$(B_n f)(x) = \sum_{k=0}^{n} \frac{n!}{(n-k)!k!} f\left(\frac{h}{n}\right) x^k (1-x)^{n-k}, \quad 0 \leqslant x \leqslant 1.$$

---

[12] 冠名源自:

H. LEBESGUE [1909]: Sur les intégrales singulières. Annales de la Faculté des Sciences de l'Université de Toulouse **1**, 25 – 117.

当 $n \to \infty$ 时 Lebesgue 常数 $\|L_n\|$ 的渐近性态, 已经有不少细致的讨论; 可参阅例如:

L. BRUTMAN [1997]: Lebesgue functions for polynomial interpolations – a survey. Annals of Numerical Mathematics **4**, 111 – 127.

A. EISINBERG; G. FEDELE; G. FRANZÈ [2004]: Lebesgue constant for Lagrange interpolation on equidistant nodes. Analysis in Theory and Applications **20**, 323 – 331.

S.J. SMITH [2006]: Lebesgue constants in polynomial interpolation. Annales Mathematicae et Informaticae **33**, 109 – 123.

R.B. PLATTE; L.N. TREFETHEN; A.B.J. KUIJLAARS [2011]: Impossibility of fast stable approximation of analytic functions from equispaced samples. SIAM Review **53**, 308 – 318.

[13] S. LOZINSKI [1948]: On a class of linear operators. Doklady Akademii Nauk SSSR **61**, 193 – 196 (俄文). 证明可在 CHENEY [1966] 的第 6 章第 5 节找到.

我们证明了相应的 Bernstein 算子 $B_n : \mathcal{C}[0,1] \to \mathcal{P}_n[0,1] \subset \mathcal{C}[0,1]$ 满足 (同上)

$$\lim_{n\to\infty} \|B_n f - f\| = 0, \quad 对每个\ f \in \mathcal{C}[0,1].$$

所以根据定理 5.4-3, 上述四条性质中至少有一条不成立. 因为每一个算子 B_n 都是线性而且连续的 (对所有 $n \geqslant 2$, $\|B_n\| = 1$), 这就得到 B_n 不是作用在所有阶数 $\leqslant n$ 的多项式上都保持不变 (见习题 2.13-1, 它在这方面提供了一个例证), 即使 B_n 的值域是空间 $\mathcal{P}_n[0,1]$ (见习题 5.4-5).

其次考虑如下定义的映射 $A_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$, 对任意 $n \geqslant 0$,

$$A_n f \in \mathcal{P}_n[0,1] \quad 且 \quad \|f - A_n f\| = \inf_{p\in\mathcal{P}_n[0,1]} \|f - p\|, \quad 对每个\ f \in \mathcal{C}[0,1]$$

(要确立 $A_n f$ 的存在比较简单, 但确立其唯一性就不大容易了; 见习题 5.4-6). 这样, 每一个映射 $A_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$, $n \geqslant 0$, 都是连续的 (见习题 5.4-6) 而且保持空间 $\mathcal{P}_n[0,1]$ 不变, 即对所有 $p \in \mathcal{P}_n[0,1]$ 成立 $A_n p = p$. 此外, 显然 Weierstrass 逼近定理 (定理 2.13-3) 意味着

$$\lim_{n\to\infty} \|A_n f - f\| = 0, \quad 对每个\ f \in \mathcal{C}[0,1].$$

因为上述四条性质中至少要有一条不成立, 仍旧利用定理 5.4-3, 剩下的只有线性作为那条不成立性质的候选者. 实际上, 每一映射 $A_n (n \geqslant 0)$ 都是非线性的 (见习题 5.4-6).

作为对比, 我们考察对任意 $n \geqslant 0$ 如下定义的映射 $P_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$:

$$P_n f \in \mathcal{P}_n[0,1] \quad 且 \quad \|f - P_n f\|_{L^2(0,1)} = \inf_{p\in\mathcal{P}_n[0,1]} \|f - p\|_{L^2(0,1)}.$$

由投影定理 (定理 4.3-1), 它在此可用是因为作为有限维子空间 $\mathcal{P}_n[0,1]$ 是 $(\mathcal{C}[0,1], \|\cdot\|_{L^2(0,1)})$ 的完备子集, 每个 $P_n : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ 都是线性且连续的算子, 此外显然它保持空间 $\mathcal{P}_n[0,1]$ 不变. 进而, Weierstrass 逼近定理意味着对每个 $f \in \mathcal{C}[0,1]$ 成立

$$\lim_{n\to\infty} \|f - P_n f\|_{L^2(0,1)} = 0$$

(此因 $\inf_{p\in\mathcal{P}_n[0,1]} \|f - p\|_{L^2(0,1)} \leqslant \inf_{p\in\mathcal{P}_n[0,1]} \|f - p\|$). 因此这种类型的多项式逼近满足上述 “理想目标” 的所有四条性质.

当然这与 Kharshiladze-Lozinski 定理的结论并不矛盾, 该定理适用于装备以 sup 范数 $\|\cdot\|$ 的空间 $\mathcal{C}[0,1]$.

习题

5.4-1 在此问题中, $L_n f$ 表示如定理 5.4-1 中所定义的函数 $f \in \mathcal{C}[a,b]$ 的 Lagrange 插值多项式.

(1) 假设 $f \in \mathcal{C}^{n+1}[a,b]$. 证明对任意 $x \in [a,b]$, 存在一点 $\xi_x \in]a,b[$ 使得

$$f(x) - L_n f(x) = \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{j=0}^{n} (x - x_j),$$

所以

$$\|L_n f - f\| \leqslant \frac{1}{(n+1)!} \|f^{(n+1)}\| \sup_{x \in [a,b]} \left| \prod_{j=0}^{n} (x - x_j) \right|.$$

注 公式

$$f(x) = L_n f(x) + \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) \prod_{j=0}^{n} (x - x_j), \quad a \leqslant x \leqslant b$$

是一维多点 Taylor 公式的一个例子. 在 7.11 节我们会看到多点 Taylor 公式在 $\mathbb{R}^n$ 中仍然成立.

提示: 给定任意点 $y \in [a,b]$ 使 $y \neq x_i$, $0 \leqslant i \leqslant n$. 对由下式定义的辅助函数 $h_y \in \mathcal{C}^{n+1}[a,b]$ 应用 Rolle 定理:

$$h_y(x) := (f - L_n f)(x) - (f - L_n f)(y) \prod_{j=0}^{n} \left(\frac{x - x_j}{y - x_j} \right).$$

(2) 证明, 如果 $f \in \mathcal{C}^{\infty}[a,b]$ 且存在一常数 C 使得 $\|f^{(n)}\| \leqslant C^n$ 对所有 $n \geqslant 0$ 成立, 则 $\lim_{n\to\infty} \|L_n f - f\| = 0$.

(3) 证明, 如果函数 f 能延拓为 $\mathbb{C}$ 的一个开子集中的解析函数, 而这个开子集包含闭集 $\{z \in \mathbb{C} : \mathrm{dist}(z; [a,b]) \leqslant 1\}$, 则 $\lim_{n\to\infty} \|L_n f - f\| = 0$.

5.4-2 (Hermite 插值的例子) 对每个整数 $n \geqslant 0$, 设给定 $n+1$ 个相异的点 $a \leqslant x_0 < x_1 < \cdots < x_n \leqslant b$.

(1) 证明给定任意函数 $f \in \mathcal{C}^1[a,b]$, 存在唯一的多项式 $p_n = p_n(f) \in \mathcal{P}_{2n+1}[a,b]$ 使得

$$p_n(x_i) = f(x_i) \quad 且 \quad p_n'(x_i) = f'(x_i), \quad 0 \leqslant i \leqslant n.$$

(2) 设空间 $\mathcal{C}^1[a,b]$ 装备以范数 $f \to \max\{\|f\|, \|f'\|\}$. 证明在 (1) 中定义的映射 $f \in \mathcal{C}^1[a,b] \to p_n(f) \in \mathcal{C}^1[a,b]$ 是线性连续的, 而且对所有 $f \in \mathcal{P}_{2n+1}[a,b]$ 成立 $p_n(f) = f$.

(3) 假设 $f \in \mathcal{C}^{2n+2}[a,b]$. 证明

$$\|p_n(f) - f\| \leqslant \frac{1}{(2n+2)!} \|f^{(2n+2)}\| \sup_{a \leqslant x \leqslant b} \left| \prod_{j=0}^{n} (x - x_j)^2 \right|.$$

提示: 利用类似于在习题 5.4-1 (1) 中所建议的推导.

5.4-3 (Hermite 插值的例子) 设给定一整数 $n \geqslant 1$ 及一个紧区间 $[a,b]$, $a < b$.

(1) 证明, 给定任意函数 $f \in \mathcal{C}^n[a,b]$, 存在唯一的多项式 $p_n = p_n(f) \in \mathcal{P}_{2n+1}[a,b]$ 使得

$$p_n^{(k)}(a) = f^{(k)}(a) \quad 及 \quad p_n^{(k)}(b) = f^{(k)}(b), \quad 0 \leqslant k \leqslant n.$$

(2) 设空间 $\mathcal{C}^n[a,b]$ 装备以范数 $f \to \max_{0\leqslant k\leqslant n}\|f^{(k)}\|$. 证明在 (1) 中定义的映射 $f\in\mathcal{C}^n[a,b]\to p_n(f)\in\mathcal{C}^n[a,b]$ 是线性连续的, 而且对所有 $f\in\mathcal{P}_{2n+1}[a,b]$ 成立 $p_n(f)=f$.

(3) 假设 $f\in\mathcal{C}^{2n+2}[a,b]$. 证明

$$\|(p_n(f)-f)^{(k)}\|\leqslant\frac{(b-a)^k}{k!(2n+2-2k)!}\|f^{(2n+2)}\|((x-a)(b-x))^{n+1-k}$$

对所有 $x\in[a,b]$ 及 $0\leqslant k\leqslant n+1$ 成立.

提示: 给定任意点 $y\in]a,b[$ 及任意的 $0\leqslant k\leqslant n+1$, 对由下式定义的辅助函数 $h_y^k\in\mathcal{C}^{2n+1-k}[a,b]$ 应用 Rolle 定理:

$$h_y^k(x)=(f-p_n(f))^{(k)}(x)-(f-p_n(f))^{(k)}(y)\left(\frac{(x-a)(b-x)}{(y-a)(b-y)}\right)^{n+1-k},\quad a\leqslant x\leqslant b.$$

5.4-4　对每个整数 $n\geqslant 1$, 设给定 $n+1$ 个相异的点 $0\leqslant x_0^n<x_1^n<\cdots<x_n^n\leqslant 1$, 又置

$$p_j^n(x):=\prod_{\begin{cases}i=0\\ i\neq j\end{cases}}^{n}\left(\frac{x-x_i^n}{x_j^n-x_i^n}\right),\quad 0\leqslant x\leqslant 1.$$

证明存在一个常数 $C>0$, 它依赖于点 x_j^n, $1\leqslant j\leqslant n$, $n\geqslant 0$, 使得

$$\sup_{0\leqslant x\leqslant 1}\left(\sum_{j=0}^{n}|p_j^n(x)|\right)\geqslant C\log n,\quad 对所有\ n\geqslant 1.$$

由此得 Lebesgue 常数 $\|L_n\|=\sup_{0\leqslant x\leqslant 1}\left(\sum_{j=0}^{n}|p_j^n(x)|\right)$ 满足 $\lim_{n\to\infty}\|L_n\|=\infty$.

5.4-5　证明, 对每个整数 $n\geqslant 1$, $\mathcal{C}[0,1]$ 在 Bernstein 算子 B_n 映射下的像是空间 $\mathcal{P}_n[0,1]$.

5.4-6　(1) 给定任意函数 $f\in\mathcal{C}[0,1]$, 证明对每个 $n\geqslant 0$, 存在唯一的多项式 $A_nf\in\mathcal{P}_n[0,1]$, 使得

$$\|f-A_nf\|=\inf_{p\in\mathcal{P}_n[0,1]}\|f-p\|.$$

因此对所有 $p\in\mathcal{P}_n[0,1]$ 成立 $A_np=p$.

(2) 设对每个 $n\geqslant 0$, $A_n:\mathcal{C}[0,1]\to\mathcal{C}[0,1]$ 表示在 (1) 中定义的映射. 证明, 对每个 $f\in\mathcal{C}[0,1]$, 存在一个常数 $C(f,n)$ 使得

$$\|A_nf-A_n\widetilde{f}\|\leqslant C(f,n)\|f-\widetilde{f}\|,\quad 对所有\ \widetilde{f}\in\mathcal{C}[0,1].$$

因此每个映射 A_n 都是连续的.

(3) 证明, 对每个 $n\geqslant 0$, 映射 $A_n:\mathcal{C}[0,1]\to\mathcal{C}[0,1]$ 都是非线性的.

提示: 这个问题绝不是平凡的. (1) 中 A_nf 的存在性可由一个简单的紧性推导得到 (如习题 2.7-1(1) 中那样), 但 A_nf 的唯一性依赖下述出色的 (但其证明不简单) de la Vallée Poussin 两样性定理[14]: 多项式 $p_n\in\mathcal{P}_n[0,1]$ 使得 $\|f-p_n\|=\inf_{p\in\mathcal{P}_n[0,1]}\|f-p\|$ 成立的充要条件是

---

[14] C.J. de la Vallée Poussin [1910]: Sur les polynômes d'approximation et la représentation approchée d'un angle. Académie Royale de Belgique, Bulletins de la Classe des Sciences **12**.

存在至少 $n+2$ 个相异的点 $0 \leqslant x_0 < x_1 < \cdots < x_{n+1} \leqslant 1$ 使得

$$|f(x_i) - p_n(x_i)| = \|f - p_n\|, \quad 0 \leqslant i \leqslant n+1,$$
$$\text{sgn}(f(x_i) - p_n(x_i)) = -\text{sgn}(f(x_{i+1}) - f(x_i)), \quad 0 \leqslant i \leqslant n.$$

实际上, 这个两样性定理不仅对子空间 $\mathcal{P}_n[0,1]$ 成立, 并且对 $\mathcal{C}[0,1]$ 的任一满足 Haar 条件[15] 的子空间 $V_n = \text{Span}(e_i)_{i=0}^n$ 成立. 给定任意的 $n+1$ 个相异的点 $x_j \in [0,1]$, $0 \leqslant j \leqslant n$, 如果 $\det(e_i(x_j)) \neq 0$, 则这个条件满足. 如果 $V_n = \mathcal{P}_n[0,1]$, $\det(e_i(x_j))$ 就是众所周知的 Vandermonde 行列式, 所以空间 $\mathcal{P}_n[0,1]$ 满足 Haar 条件[16].

5.5 Banach-Steinhaus 定理的应用: Fourier 级数的发散

我们回忆一下, $\mathcal{C}_{\text{per}}[0,2\pi]$ 表示所有连续的 2π 周期函数 $g:\mathbb{R}\to\mathbb{R}$ 组成的 Banach 空间, 其中装备以由 $\|g\| = \sup_{0\leqslant\theta\leqslant 2\pi}|g(\theta)|$ 定义的 sup 范数 $\|\cdot\|$; $\mathcal{Q}_n[0,2\pi]$ 表示所有阶数 $\leqslant n$ 的实 2π 周期三角多项式构成的空间, g 的第 n 个 Fourier 部分和由下式定义:

$$(S_n g)(\theta) := \frac{a_0}{2} + \sum_{k=1}^{n} a_k \cos k\theta + \sum_{k=1}^{n} b_k \sin k\theta, \quad 0 \leqslant \theta \leqslant 2\pi,$$

其中

$$a_k := \frac{1}{\pi}\int_0^{2\pi} g(\theta)\cos k\theta d\theta, \quad k \geqslant 0,$$
$$b_k := \frac{1}{\pi}\int_0^{2\pi} g(\theta)\sin k\theta d\theta, \quad k \geqslant 1.$$

我们后来还证明了 (定理 2.14-2) 如下定义的 Fejér 算子 $F_n : \mathcal{C}_{\text{per}}[0,2\pi] \to \mathcal{Q}_{n-1}[0,2\pi] \subset \mathcal{C}_{\text{per}}[0,2\pi]$:

$$F_n : g \in \mathcal{C}_{\text{per}}[0,2\pi] \to F_n g := \frac{1}{n}(S_0 g + S_1 g + \cdots + S_{n-1} g), \quad \text{对每个 } n \geqslant 1$$

具有以下性质: 对任意函数 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$, 成立

$$\lim_{n\to\infty} \|F_n g - g\| = 0.$$

与此形成对比的是, 我们现在证明, 存在函数 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$, 其第 n 个 Fourier 部分和 $S_n g$ 不一致收敛于 g (本节标题所示 “Fourier 级数的发散” 即是按此意义理解). 这一结果是下面定理的一个推论, 在该定理中以与上节同样的方式 (试比较定理 5.4-2) 应用 Banach-Steinhaus 定理, 证明存在函数 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$ 使得 $\sup_{n\geqslant 0}\|S_n g\| = \infty$.

---

[15] A. HAAR [1918]: Die Minkowskische Geometrie und die Annäherung an stetige Funktionen. Mathematische Annalen **78**, 294 – 311.

[16] (1) 与 (2) 的证明, 例如可参阅 CHENEY [1966], 第 3 章第 4 和 5 节.

定理 5.5-1 存在函数 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$, 其第 n 个 Fourier 部分和 $S_n g$ 满足

$$\sup_{n\geqslant 0}\|S_n g\| = \infty.$$

这一性质使得对这种函数 g, $S_n g$ 不可能一致收敛于 g.

证明 (i) 对每个 $n \geqslant 0$, 第 n 个 Fourier 部分和定义的线性算子 $S_n : \mathcal{C}_{\text{per}}[0,2\pi] \to \mathcal{C}_{\text{per}}[0,2\pi]$ 是连续的, 且其范数由下式给出:

$$\|S_n\| = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\sin\left(n+\dfrac{1}{2}\right)\varphi}{\sin\dfrac{1}{2}\varphi}\right|d\varphi.$$

设 Dirichlet 核 $D_n \in \mathcal{C}_{\text{per}}[0,2\pi]$ 由下式定义:

$$D_n(\varphi) := \frac{1}{2\pi}\frac{\sin\left(n+\dfrac{1}{2}\right)\varphi}{\sin\dfrac{1}{2}\varphi}, \quad 0 \leqslant \varphi \leqslant 2\pi.$$

则任意函数 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$ 的第 n 个 Fourier 部分和 $S_n g$ 也可以写为 (见习题 2.14-1)

$$S_n g(\theta) = \int_{-\pi}^{\pi} g(\theta+\varphi)D_n(\varphi)d\varphi,$$

因此

$$\|S_n\| \leqslant \sup_{g\neq 0}\frac{\|S_n g\|}{\|g\|} \leqslant \int_{-\pi}^{\pi}|D_n(\varphi)|d\varphi.$$

设函数 $g_n : [-\pi,\pi] \to \mathbb{R}$ 如下定义:

$$g_n(\varphi) := \operatorname{sgn} D_n(\varphi), \quad -\pi \leqslant \varphi \leqslant \pi,$$

而且对充分小的 $\varepsilon > 0$, 令 $g_n^\varepsilon : [-\pi,\pi] \to \mathbb{R}$ 表示连续的分段仿射函数, 它在 $[-\pi,\pi] - I_n^\varepsilon$ 上等于 g_n, 其中 I_n^ε 表示 $[-\pi,\pi]$ 与诸长度为 ε 的开区间并集的交集, 这些开区间的中心位于区间 $[-\pi,\pi]$ 上 Dirichlet 核 D_n 的零点 (图 5.5-1).

这样就有 $\|g_n^\varepsilon\| = 1$, 且

$$\|S_n g_n^\varepsilon\| \geqslant |S_n g_n^\varepsilon(0)| = \left|\int_{-\pi}^{\pi} g_n^\varepsilon(\varphi)D_n(\varphi)d\varphi\right|.$$

容易验证

$$\lim_{\varepsilon\to 0}\int_{-\pi}^{\pi} g_n^\varepsilon(\varphi)D_n(\varphi)d\varphi = \int_{-\pi}^{\pi} g_n(\varphi)D_n(\varphi)d\varphi = \int_{-\pi}^{\pi}|D_n(\varphi)|d\varphi,$$

而且

$$\|S_n\| = \sup_{g\neq 0}\frac{\|S_n g\|}{\|g\|} \geqslant \frac{\|S_n g_n^\varepsilon\|}{\|g_n^\varepsilon\|} = \|S_n g_n^\varepsilon\|,$$

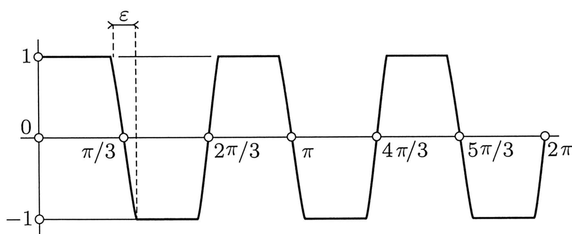


图 5.5-1 定理 5.5-1 证明中的函数 g_n^ε, 这里画出的是 $n=5$ 的情形

所以有 $\|S_n\| \geqslant \int_{-\pi}^{\pi} |D_n(\varphi)| d\varphi$. 故 $\|S_n\| = \int_{-\pi}^{\pi} |D_n(\varphi)| d\varphi$, 如所宣示.

(ii) 范数 $\|S_n\|$ 满足

$$\|S_n\| \geqslant \frac{4}{\pi^2} \log n, \quad \text{对所有 } n \geqslant 1.$$

这个不等式可由初等计算得到, 作为习题留给读者 (习题 5.5-1).

(iii) 第 n 个部分和 $S_n g$ 满足条件 $\sup_{n\geqslant 0} \|S_n g\| = \infty$ 的函数 $g \in \mathcal{C}_{\text{per}}[0, 2\pi]$ 的存在性.

如果对每一个 $g \in \mathcal{C}_{\text{per}}[0, 2\pi]$ 我们都有 $\sup_{n\geqslant 1} \|S_n g\| < \infty$, 那么 Banach-Steinhaus 定理 (定理 5.3-1) 意味着 $\sup_{n\geqslant 1} \|S_n\| < \infty$. 但这与我们可由 (ii) 得到的关系式 $\lim_{n\to\infty} \|S_n\| = \infty$ 矛盾. 所以 $\sup_{n\geqslant 0} \|S_n g\| = \infty$ 至少对一个函数 $g \in \mathcal{C}_{\text{per}}[0, 2\pi]$ 成立. □

注 如果代替 Banach-Steinhaus 定理, 我们利用它的推论 (定理 5.3-2), 我们也能得到如下结论: 存在函数 $g \in \mathcal{C}_{\text{per}}[0, 2\pi]$, 其第 n 个 Fourier 部分和 $S_n g$ 不按 sup 范数收敛于 g. 但我们得不到 $\sup_{n\geqslant 1} \|S_n g\| = \infty$ 这一结论.

在定理 5.5-1 的证明中自然出现的范数

$$\|S_n\| = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{\sin\left(n+\frac{1}{2}\right)\varphi}{\sin\frac{\varphi}{2}} \right| d\varphi, \quad n \geqslant 0$$

称为 Lebesgue 常数[17] (注意与定理 5.4-2 的证明中出现的 “另一个” Lebesgue 常数 $\|L_n\|$ 进行区分).

正如多项式插值的发散现象并不局限于 Lagrange 插值 (5.4 节), 三角多项式逼近的发散现象也并非 Fourier 级数所特有, 这可由下述美妙的结果看出 (毫不奇怪, 其证明仍依赖于 Banach-Steinhaus 定理).

---

[17] 冠名源自:

H. LEBESGUE [1909]: Sur les intégrales singulières. Annales de la Faculté des Sciences de l'Université de Toulouse **1**, 25 – 117.

定理 5.5-2 (Kharshiladze-Lozinski 三角逼近定理[18]) 对任意由映射 $B_n : \mathcal{C}_{\text{per}}[0,2\pi] \to \mathcal{Q}_n[0,2\pi] \subset \mathcal{C}_{\text{per}}[0,2\pi]$ 组成的序列 $(B_n)_{n=0}^{\infty}$, 若映射 B_n 都是线性连续的, 而且保持所有阶数 $\leqslant n$ 的三角多项式不变, 即 $B_n q = q$ 对所有 $q \in \mathcal{Q}_n[0,2\pi]$ 及所有 $n \geqslant 0$ 成立, 则必有

$$\|B_n\| \geqslant \|S_n\|, \quad \text{对所有 } n \geqslant 0,$$

其中 S_n 对每个 $n \geqslant 0$ 表示与第 n 个 Fourier 部分和相关联的算子.

因为根据 Banach-Steinhaus 定理又有

$$\sup_{n \geqslant 0} \|B_n g\| = \infty, \quad \text{对某个 } g \in \mathcal{C}_{\text{per}}[0,2\pi],$$

所以定理 5.5-2 意味着, 不可能存在任何这样的序列 $(B_n)_{n=0}^{\infty}$, 它另外还满足以下性质: $\lim_{n\to\infty} \|B_n g - g\| = 0$ 对所有的 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$ 成立.

作为比较, 考察 Fejér 算子 F_n. 它满足 $\lim_{n\to\infty} \|F_n g - g\| = 0$ 对所有 $g \in \mathcal{C}_{\text{per}}[0,2\pi]$ 成立 (定理 2.14-2), 故它必定至少缺少上述四条性质中的一条. 实际上, 它们都是线性连续的 ($\|F_n\| = 1$, 对所有 $n \geqslant 1$), 但并不保持所有阶数 $\leqslant n$ 的三角多项式不变 (习题 2.14-1(3)).

习题

5.5-1 证明 Lebesgue 常数 $\|S_n\| := \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} \left| \dfrac{\sin\left(n+\frac{1}{2}\right)\varphi}{\sin\frac{1}{2}\varphi} \right| d\varphi$ (其在定理 5.5-1 的证明中起关键作用) 满足 $\|S_n\| \geqslant \dfrac{4}{\pi^2}\log n$, 对所有 $n \geqslant 1$.

5.6 Banach 开映射定理; 首选应用: 两点边值问题的适定性

Baire 定理的另一个基本结论是下述充分条件, 它判定从一个 Banach 空间到另一个 Banach 空间的连续线性算子是开映射, 即将开集映为开集.

定理 5.6-1 (Banach 开映射定理[19]) 设 X 及 Y 是 Banach 空间, 而 $A \in \mathcal{L}(X;Y)$ 是满射.

则 X 的任一开子集 U 在映射 A 下的直接像 $A(U)$ 是 Y 的开子集.

---

[18] S. Lozinski [1948]: On a class of linear operators. Doklady Akademii Nauk SSSR **61**, 193–196 (俄文).

在 Cheney [1996] 第 6 章第 5 节中可找到该定理的证明.

[19] S. Banach [1932]: Théorie des Opérateurs Linéaires. Monngrafje Matematyczne, Warsaw.

证明 在整个证明中, 下述符号分别用来表示空间 X 中, 球心位于 X 的原点的开球及 Y 中的开球:

$$B_r := \{x \in X; \|x\| < r\} \quad 及 \quad B(y; s) = \{\widetilde{y} \in Y; \|\widetilde{y} - y\| < s\}.$$

另外, 给定一个向量空间 Z, 向量 $z_0 \in Z$, 数量 α 及一个子集 $A \subset Z$. 我们设

$$\{z_0\} + A := \{(z_0 + z) \in Z; z \in A\} \quad 及 \quad \alpha A := \{(\alpha z) \in Z; z \in A\}.$$

故如果 $z_0 \in A$, 则 $\{z_0\} + A \subset 2A$ 且 A 是凸的.

(i) 集合 $\overline{A(B_1)}$ 包含一个开球.

给定任意的 $y \in Y$, 因为 A 是满射, 存在 $x \in X$ 使得 $y = Ax$ (这是唯一用到满射假定的地方). 由于 $x \in B_n$ 对某个整数 $n \geqslant 1$ 成立, 这说明 $Y = \bigcup_{n=1}^{\infty} A(B_n)$; 那就更有

$$Y = \bigcup_{n=1}^{\infty} \overline{A(B_n)}.$$

空间 Y 是完备的, Baire 定理 (这里用的是其推论形式, 定理 5.1-3(b)) 说明存在一个整数 $n_0 \geqslant 1$ 使得

$$\operatorname{int} \overline{A(B_{n_0})} \neq \varnothing.$$

A 是线性的, 故 $\overline{A(B_1)} = \dfrac{1}{n_0}\overline{A(B_{n_0})}$. 所以 $\operatorname{int} \overline{A(B_1)} \neq \varnothing$, 集合 $\overline{A(B_1)}$ 包含一个开球.

(ii) 集合 $\overline{A(B_1)}$ 包含一个球心在原点 Y 的开球.

由 (i), 存在 $y \in Y$ 及 $s > 0$ 使得 $B(y; 2s) \subset \overline{A(B_1)}$, 因此

$$B(0; 2s) = \{-y\} + B(y; 2s) \subset \{-y\} + \overline{A(B_1)}.$$

由于 $-y \in \overline{A(B_1)}$ (此因 $y \in \overline{A(B_1)}$ 而 A 是线性的), 故有

$$\{-y\} + \overline{A(B_1)} \subset 2\overline{A(B_1)}.$$

由得到的包含关系 $B(0, 2s) \subset 2\overline{A(B_1)}$, 再考虑到 A 是线性的, 就有

$$B(0; s) \subset \overline{A(B_1)}.$$

(iii) 集合 $A(B_1)$ 包含一个球心在 Y 的原点的开球.

为了证明这一断言, 我们将证明 $B\left(0; \dfrac{s}{2}\right) \subset A(B_1)$, 其中 $s > 0$ 是出现在上述 (ii) 部分中的球 $B(0; s)$ 的半径. 这意味着, 给定任意的 $y \in B\left(0; \dfrac{s}{2}\right)$, 我们要找 $x \in B_1$ 使得 $y = Ax$. 设 $y \in B\left(0; \dfrac{s}{2}\right)$ 是给定的.

因为由 (ii) 有 $y \in B\left(0; \dfrac{s}{2}\right) \subset \overline{A(B_{1/2})}$, 故存在 $x_1 \in B_{1/2}$ 使得 $\|y - Ax_1\| < \dfrac{s}{2^2}$; 还是由 (ii) 有

$$(y - Ax_1) \in B\left(0; \frac{s}{2^2}\right) \subset \overline{A(B_{1/2^2})},$$

故存在 $x_2 \in B_{1/2^2}$ 使得 $\|y - Ax_1 - Ax_2\| < \dfrac{s}{2^3}$, 如此等等. 按此方式, 我们构造出点 $x_n \in X$ 的序列 $(x_n)_{n=1}^{\infty}$, 它具有以下性质:

$$x_n \in B_{1/2^n} \quad 且 \quad \left\| y - A\left(\sum_{k=1}^{n} x_k\right) \right\| < \frac{s}{2^{n+1}}, \quad 对所有\ n \geqslant 1.$$

因为级数 $\sum_{n=1}^{\infty} x_n$ 是一致收敛的 (此因 $\|x_n\| < \dfrac{1}{2^n}$ 对每个 $n \geqslant 1$ 成立) 且空间 X 是完备的, 所以存在 $x \in X$ 使得

$$\lim_{n\to\infty} \sum_{k=1}^{n} x_k = x \quad 且 \quad \|x\| \leqslant \sum_{k=1}^{\infty} \|x_k\| < \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n+1}} + \cdots = 1$$

(见定理 3.6-1), 故 $x \in B_1$. 进而因 A 是连续的,

$$y = \lim_{n\to\infty} A\left(\sum_{k=1}^{n} x_k\right) = Ax.$$

本段的断言得到证明.

(iv) 映射 A 是开的.

给定 X 的任一开子集 U 及给定任意点 $y \in A(U)$, 我们必须找到 $\sigma > 0$ 使 $B(y;\sigma) \subset A(U)$. 为此设 $x \in U$ 使 $y = Ax$.

因 U 是开的, 所以存在 $r > 0$ 使得 $B(x;r) = \{x\} + B_r \subset U$. 又由 (iii), 存在 $\sigma > 0$ 使得 $B(0;\sigma) \subset A(B_r)$. 因此有

$$B(y;\sigma) = \{y\} + B(0;\sigma) \subset \{y\} + A(B_r) = A(\{x\} + B_r) \subset A(U). \qquad \square$$

下述由 Banach 开映射定理易得的结论, 以后将称为 Banach 开映射定理的推论, 常被用作判定线性算子逆的连续性的充分条件.

定理 5.6-2 (Banach 开映射定理的推论) 设 X 与 Y 是 Banach 空间, 而 $A \in \mathcal{L}(X;Y)$ 是双射. 则 $A^{-1} \in \mathcal{L}(X;Y)$.

证明　X 与 Y 中的开球仍用定理 5.6-1 证明中的符号表示. 因为映射 $A^{-1}: Y \to X$ 也是线性的 (定理 2.9-1(b)), 故只需证明它在原点连续 (定理 2.9-2(b)). 根据在一点连续的定义 (1.11 节), 只要证明给定任意开球 $B_r \subset X$, 存在一个开球 $B(0;\sigma) \subset Y$, 使得

$$A^{-1}(B(0;\sigma)) \subset B_r.$$

但由于 A 是双射, 包含关系 $A^{-1}(B(0;\sigma)) \subset B_r$ 等价于包含关系

$$B(0;\sigma) \subset A(B_r).$$

而 Banach 开映射定理恰好证明了上述包含关系对某个 $\sigma > 0$ 成立. $\square$

下面以两点边值问题的应用作为第一个实例, 显示 Banach 开映射定理推论的威力. 只要假设对所有的右端 $f \in \mathcal{C}[0,1]$, 解 $u \in \mathcal{C}^2[0,1]$ 存在且是唯一的, 由这个定理就可以证明此问题在下述意义下是适定的: f 在 sup 范数下 "小" 的扰动导致 u, u' 及 u'' 也在 sup 范数下 "小" 的变化. 值得关注的是, 如此强有力的连续性结果可以在很少的假设下得到. 例如, 如果 $a(x) = -1$ 而 $\|b\| + \|c\|$ 充分小 (习题 3.9-1), 或者如果 $a(x) = -1, b(x) = 0, c(x) \geqslant 0, 0 \leqslant x \leqslant 1$ (习题 3.10-3; 实际这个不等式可放宽为 $c(x) \geqslant \gamma > -\pi^2, 0 \leqslant x \leqslant 1$; 见习题 9.14-3) 就满足这些假设.

定理 5.6-3 设给定函数 $a, b, c \in \mathcal{C}[0,1]$ 使得两点边值问题

$$a(x)u''(x) + b(x)u'(x) + c(x)u(x) = f(x), \quad 0 \leqslant x \leqslant 1, \quad u(0) = u(1) = 0$$

对每个 $f \in \mathcal{C}[0,1]$ 有一个且只有一个解 $u \in \mathcal{C}^2[0,1]$. 则存在常数 C 使得

$$\|u\| + \|u'\| + \|u''\| \leqslant C\|f\|, \quad 对所有\ f \in \mathcal{C}[0,1],$$

其中 $\|\cdot\|$ 表示空间 $\mathcal{C}[0,1]$ 的 sup 范数.

证明 空间

$$X := \{v \in \mathcal{C}^2[0,1]; v(0) = v(1) = 0\}$$

装备范数 $v \to (\|v\| + \|v'\| + \|v''\|)$ 后是 Banach 空间 (见习题 5.6-2), 而线性算子 $L : v \in X \to Lv \in Y := \mathcal{C}[0,1]$, 其中

$$Lv(x) := a(x)v''(x) + b(x)v'(x) + c(x)v(x), \quad 0 \leqslant x \leqslant 1$$

是连续的, 此因

$$\|Lv\| \leqslant \max\{\|a\|, \|b\|, \|c\|\}(\|v\| + \|v'\| + \|v''\|), \quad 对所有\ v \in X.$$

这样, 定理的结论可立即由 Banach 开映射定理的推论 (定理 5.6-2) 得到. □

Banach 开映射定理的另一个推论是下述无穷维空间两个范数等价的充分条件. 特别地, 它将为 Banach 闭图像定理 (定理 5.7-1) 提供一个简洁的证明.

定理 5.6-4 设 $\|\cdot\|$ 与 $\|\cdot\|'$ 是同一个向量空间的两个具有以下性质的范数: 空间 $(X, \|\cdot\|)$ 与 $(X, \|\cdot\|')$ 都是完备的, 且存在常数 C 使得

$$\|x\|' \leqslant C\|x\|, \quad 对所有\ x \in X.$$

则两个范数 $\|\cdot\|$ 与 $\|\cdot\|'$ 是等价的.

证明 由假设, 双射且线性的恒等映射 $\iota : (X, \|\cdot\|) \to (X, \|\cdot\|')$ 是连续的. 所以定理 5.6-2 说明其逆映射 $\iota^{-1} : (X, \|\cdot\|') \to (X, \|\cdot\|)$ 也是连续的. 这意味着存在常数 C' 使得 $\|x\| \leqslant C'\|x\|'$ 对所有 $x \in X$ 成立. 因此两个范数是等价的. □

Banach 开映射定理的另一个有趣的应用在习题 5.6-3 中给出.

习题

5.6-1 是否存在向量空间 X 及其上的两个范数 $\|\cdot\|$ 与 $\|\cdot\|'$ 使得空间 $(X,\|\cdot\|)$ 与 $(X,\|\cdot\|')$ 都是完备的, 但两个范数 $\|\cdot\|$ 与 $\|\cdot\|'$ 不等价?

5.6-2 在这个问题中 $\|\cdot\|$ 表示空间 $\mathcal{C}[0,1]$ 的 sup 范数, 而 $m \geqslant 1$ 是整数.

(1) 证明函数 $v \in \mathcal{C}^m[0,1] \to (\|v\| + \|v'\| + \cdots + \|v^{(m)}\|)$ 定义空间 $\mathcal{C}^m[0,1]$ 上的一个范数, 且使该空间为 Banach 空间.

(2) 证明空间 $(\mathcal{C}^m[0,1], \|\cdot\|)$ 是不完备的.

提示: 不要去展示一个不收敛的 Cauchy 序列, 直接利用 (1) 在 $m=1$ 的情况以及定理 5.6-4.

5.6-3 设 X 及 Y 是 Banach 空间. 利用 Banach 开映射定理证明集合 $\{A \in \mathcal{L}(X;Y); A$ 是满射$\}$ 在装配算子范数的空间 $\mathcal{L}(X;Y)$ 中是开的.

换言之, 如果 $A \in \mathcal{L}(X;Y)$ 使得方程 $Ax = y$ 对任意的 $y \in Y$ 有至少一个解 $x \in X$, 则方程 $\widetilde{A}x = y$ 对任意的 $y \in Y$ 也至少有一个解 $x \in X$, 如果 $\widetilde{A}$ 充分接近 A.

5.7 Banach 闭图像定理; 首选应用: Hellinger-Toeplitz 定理

给定两个集合 X 与 Y, 映射 $A: X \to Y$ 的图像 Gr A 是如下定义的乘积 $X \times Y$ 的子集:

$$\text{Gr } A := \{(x, Ax) \in X \times Y; x \in X\}.$$

若 X 与 Y 是拓扑空间, 一个映射 $A: X \to Y$ 被称为闭的, 如果其图像 GrA 在积空间 $X \times Y$ (装备以积拓扑, 见 1.6 节) 中是闭的.

所以, 如果 X 与 Y 是距离空间, 映射 $A: X \to Y$ 是闭的充分必要条件是

$$\lim_{n\to\infty} x_n = x \text{ 在 } X \text{ 中}, \quad \lim_{n\to\infty} Ax_n = y \text{ 在 } Y \text{ 中}$$

意味着 $y = Ax$ (见定理 1.11-1). 任何连续映射 $A: X \to Y$ 有闭图像 (因为 $\lim_{n\to\infty} x_n = x$ 在 X 中成立意味着 $\lim_{n\to\infty} Ax_n = Ax$ 而且在距离空间中一个收敛序列的极限是唯一的). 然而反之一般不一定成立: 容易构造赋范向量空间之间的闭线性算子的简单实例, 但它们却是不连续的 (这种例子可参见习题 5.7-1).

但值得注意的是, 如果 X 与 Y 都是 Banach 空间, Banach 开映射定理的一个简单的推论说明, 任何闭的线性映射 $A: X \to Y$ 都是连续的.

定理 5.7-1 (Banach 闭图像定理[20]) 设 X 与 Y 是 Banach 空间, 而 $A: X \to Y$ 是闭线性算子. 则 $A \in \mathcal{L}(X;Y)$.

[20] S. Banach [1932]: Théorie des Opérateurs Linéaires. Monografje Matematyczne, Volume 1, Warsaw.

证明 在 X 上定义另一个范数

$$\|x\|' := \|x\|_X + \|Ax\|_Y, \quad \text{对所有 } x \in X.$$

由范数 $\|\cdot\|'$ 的定义知, 对任何关于范数 $\|\cdot\|'$ 的 Cauchy 序列 $(x_n)_{n=1}^{\infty}$, $(x_n)_{n=1}^{\infty}$ 及 $(Ax_n)_{n=1}^{\infty}$ 分别是空间 X 及 Y 中的 Cauchy 序列. 由于两个空间都是完备的, 故存在 $x \in X$ 及 $y \in Y$ 使得

$$\lim_{n\to\infty} x_n = x \text{ 在 } X \text{ 中}, \quad \lim_{n\to\infty} Ax_n = y \text{ 在 } Y \text{ 中}.$$

由假设, A 是闭的, 所以 $y = Ax$. 因此, 当 $n \to \infty$ 时,

$$\|x_n - x\|' = \|x_n - x\|_X + \|Ax_n - Ax\|_Y = \|x_n - x\|_X + \|Ax_n - y\|_Y \to 0.$$

这说明 $(X, \|\cdot\|')$ 也是完备的.

因为 $\|x\| \leqslant \|x\|'$ 对所有 $x \in X$ 成立, 定理 5.6-4 (其本身就是 Banach 开映射定理的一个推论) 说明, 存在常数 C 使得

$$\|Ax\| \leqslant \|x\| + \|Ax\| = \|x\|' \leqslant C\|x\|, \quad \text{对所有 } x \in X.$$

所以线性算子 A 是连续的. □

下面引人瞩目的结果 ("引人瞩目" 在于这么强的一个结论竟然是从一个微不足道的假设推得的) 构成了 Banach 闭图像定理的第一个应用 (其他应用在习题 5.7-2 及 5.7-3 中给出).

定理 5.7-2 (Hellinger-Toeplitz 定理[21]) 设 $(X, (\cdot,\cdot))$ 是 Hilbert 空间, 而 $A: X \to X$ 是自伴线性算子 (4.10 节), 它满足

$$(Ax, y) = (x, Ay), \quad \text{对所有 } x, y \in X.$$

则 A 是连续的.

证明 由 Banach 闭图像定理, 只需证明 A 是闭的. 设 $(x_n)_{n=1}^{\infty}$ 是元素 $x_n \in X$ 的一个序列, 使得当 $n \to \infty$ 时, $x_n \to x \in X$ 而 $Ax_n \to y \in X$. 根据内积的连续性 (定理 4.1-1), 对任意的 $z \in X$ 当 $n \to \infty$ 时有

$$(Ax_n, z) \to (y, z) \quad \text{及} \quad (Ax_n, z) = (x_n, Az) \to (x, Az) = (Ax, z).$$

因此, $(y, z) = (Ax, z)$ 对所有 $z \in X$ 成立. 这就有 $y = Ax$, 说明 A 是闭的. Banach 闭图像定理意味着 A 是连续的 (空间 X 是完备的). □

[21] E. HELLINGER; O. TOEPLITZ [1910]: Grundlagen für eine Theorie der unendichen Matrizen. Mathematische Annalen **69**, 281–330.

注　(1) 如果 X 只是一个内积空间, 上述证明说明线性算子 A 是闭的.

(2) 容易看出, 任一映射 $A: X \to X$ 如果满足 $(Ax, y) = (x, Ay)$ 对所有 $x, y \in X$ 成立, 则肯定是线性的.

我们回忆一下, 把 $\mathbb{K} = \mathbb{R}$ 上的 Hilbert 空间中的自伴线性算子称为对称的. 这种对称算子的例子来自线性椭圆边值问题的弱形式, 是下一章的中心课题.

习题

5.7-1　在本题中, $\|\cdot\|$ 表示空间 $\mathcal{C}[0,1]$ 的 sup 范数.

(1) 证明由 $(Av)(x) = v'(x), 0 \leqslant x \leqslant 1$, 对所有 $v \in \mathcal{C}^1[0,1]$ 定义的线性算子 $A: (\mathcal{C}^1[0,1]; \|\cdot\|) \to (\mathcal{C}[0,1], \|\cdot\|)$ 是不连续的.

(2) 证明 A 是闭的 (闭图像定理的一个应用, 它再次证明了 $(\mathcal{C}^1[0,1]; \|\cdot\|)$ 不是 Banach 空间; 见习题 5.6-2(2)).

5.7-2　给定 $1 < p < \infty$, 设 $q > 1$ 由 $\dfrac{1}{p} + \dfrac{1}{q} = 1$ 确定. 则给定任意的 $a = (a_i)_{i=1}^{\infty} \in \ell^q$, 级数 $\sum_{i=1}^{\infty} a_i x_i$ 对所有 $x = (x_i)_{i=1}^{\infty} \in \ell^p$ 在 $\mathbb{K}$ 中收敛, 因此由 Hölder 不等式 (定理 2.4-1) 有

$$\left|\sum_{i=1}^{\infty} a_i x_i\right| \leqslant \|a\|_q \|x\|_p, \quad \text{对所有 } x = (x_i)_{i=1}^{\infty} \in \ell^p.$$

反之, 证明如果数 a_i 的序列 $a = (a_i)_{i=1}^{\infty}$ 使得级数 $\sum_{i=1}^{\infty} a_i x_i$ 对所有 $(x_i)_{i=1}^{\infty} \in \ell^p$ 都收敛, 则 $a \in \ell^q$.

提示: 证明由 $A: x = (x_i)_{i=1}^{\infty} \in \ell^p \to Ax = (\sum_{i=1}^{j} a_i x_i)_{j=1}^{\infty}$ 定义的线性算子 $A: \ell^p \to \ell^{\infty}$ 是闭的, 再应用闭图像定理.

5.7-3　给定 $1 < p < \infty$, 设 $q > 1$ 由 $\dfrac{1}{p} + \dfrac{1}{q} = 1$ 确定. 又设 $(a_{ij})_{i,j=1}^{\infty}$ 是具有下述性质的 “无穷矩阵”: 给定任意的 $x = (x_j)_{j=1}^{\infty} \in \ell^p$, 每个级数 $\sum_{j=1}^{\infty} a_{ij} x_j, i \geqslant 1$, 都收敛. 此外, $y = (y_i)_{i=1}^{\infty} \in \ell^q$, 其中 $y_i := \sum_{j=1}^{\infty} a_{ij} x_j, i \geqslant 1$.

证明由 $Ax := y$ 对所有 $x \in \ell^p$ 定义的线性算子 $A: \ell^p \to \ell^q$ 是连续的.

提示: 利用习题 5.7-2 证明 $A: \ell^p \to \ell^q$ 是闭的, 再应用闭图像定理.

5.8　向量空间中的 Hahn-Banach 定理

向量空间中的 Hahn-Banach 定理 (下面的定理 5.8-1) 是线性泛函分析中的两个核心定理之一, 另一个是 Baire 定理 (定理 5.1-2). 而且这个定理常被称为 Hahn-Banach 定理的解析形式, 其要旨将遍及本章的其余部分. 这主要体现在下两节中证明的推论上. 在这些推论中, 基于其重要性最为突出的是这两个: 赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1) 及 Hahn-Banach 定理的几何形式 (定理 5.10-1 及 5.10-2).

在此, 我们只给出这些定理在实情况下的证明, 因为在本书的其余部分只会碰到这种情况. 对于复情况的证明, 其假设要受更多的限制, 作为习题留给读者 (见习题 5.8-1).

注意, Hahn-Banach 定理的证明需要选择公理 (利用 Zorn 引理), 而无须用 Baire 定理.

定理 5.8-1 (实向量空间中的 Hahn-Banach 定理[22]) 设 X 是实向量空间, 而 p 是 X 上的次线性泛函, 即一个满足下述性质的函数 $p : X \to \mathbb{R}$:

$$p(\alpha x) = \alpha p(x), \quad \text{对所有 } \alpha > 0 \text{ 及所有 } x \in X,$$
$$p(x+y) \leqslant p(x) + p(y), \quad \text{对所有 } x, y \in X.$$

另设 Y 是 X 的子空间, $\ell : Y \to \mathbb{R}$ 是 Y 上的线性泛函, 它满足

$$\ell(y) \leqslant p(y), \quad \text{对所有 } y \in Y.$$

则存在线性泛函 $\widetilde{\ell} : X \to \mathbb{R}$, 它满足

$$\widetilde{\ell}(y) = \ell(y), \quad \text{对所有 } y \in Y. \quad \text{而 } \widetilde{\ell}(x) \leqslant p(x), \quad \text{对所有 } x \in X.$$

证明 (i) 假定 $Y \subsetneqq X$, 选取任一元素 $x_0 \in X - Y$ (故 $x_0 \neq 0$), 定义 X 的子空间

$$\text{Dom}\, f := \{(\alpha x_0 + y) \in X; \alpha \in \mathbb{R}, y \in Y\},$$

它显然包含 Y. 我们现在证明存在线性泛函 $f : \text{Dom}\, f \to \mathbb{R}$, 它满足

$$f(y) = \ell(y), \quad \text{对所有 } y \in Y \quad \text{及} \quad f(x) \leqslant p(x), \quad \text{对所有 } x \in \text{Dom}\, f.$$

找 f 归结为找一实数 $\lambda := f(x_0)$ 使得

$$\begin{aligned} f(\alpha x_0 + y) &= \alpha\lambda + f(y) = \alpha\lambda + \ell(y) \\ &\leqslant p(\alpha x_0 + y), \quad \text{对所有 } \alpha \in \mathbb{R} \text{ 及所有 } y \in Y. \end{aligned}$$

因为这个不等式对 $\alpha = 0$ 成立, 故只需求 $\lambda \in \mathbb{R}$ 使其满足

$$\begin{aligned} \lambda &\leqslant \alpha^{-1}(p(\alpha x_0 + y) - \ell(y)) \\ &= p(x_0 + \alpha^{-1} y) - \ell(\alpha^{-1} y), \quad \text{对所有 } \alpha > 0 \text{ 及所有 } y \in Y \end{aligned}$$

[22] 这个结果首先 (以定理 5.9-1 中赋范向量空间形式) 发表在:

H. HAHN [1927]: Über lineare Gleichungssysteme in linearen Räumen. Journal de Crelle **157**, 214–229.

后又由 Stefan Banach 于 1929 年独立地得到.

及

$$\lambda \geqslant \alpha^{-1}(p(\alpha x_0 + y) - \ell(y)) = \alpha^{-1}(p(-\alpha(-x_0 - \alpha^{-1}y)) - \ell(y))$$
$$= -p(-x_0 - \alpha^{-1}y) + \ell(-\alpha^{-1}y), \quad \text{对所有 } \alpha < 0 \text{ 及所有 } y \in Y.$$

$\ell : Y \to \mathbb{R}$ 的线性及 $p : X \to \mathbb{R}$ 的次线性意味着对所有的 $u, v \in Y$,

$$\ell(u) + \ell(v) = \ell(u+v) \leqslant p(u+v) = p(-x_0 + u + x_0 + v)$$
$$\leqslant p(-x_0 + u) + p(x_0 + v),$$

因此

$$-p(-x_0 + u) + \ell(u) \leqslant p(x_0 + v) - \ell(v), \quad \text{对所有 } u, v \in Y.$$

又因

$$a := \sup_{u \in Y}\{-p(-x_0 + u) + \ell(u)\} \leqslant b := \inf_{v \in Y}\{p(x_0 + v) - \ell(v)\},$$

故只要选取 λ 满足 $a \leqslant \lambda \leqslant b$ 即可.

(ii) 设 $\mathcal{F}$ 表示定义在 X 的包含 Y 的子空间 $\mathrm{Dom}\, f$ 上、满足下述条件的所有线性泛函 $f : \mathrm{Dom}\, f \to \mathbb{R}$ 的集合:

$$f(y) = \ell(y), \quad \text{对所有 } y \in Y, \quad f(x) \leqslant p(x), \quad \text{对所有 } x \in \mathrm{Dom}\, f.$$

集合 $\mathcal{F}$ 非空, 此因 $\ell \in \mathcal{F}$. 此外, $\mathcal{F}$ 按关系 $\preccurlyeq$ 是半序的 (1.3 节), 其中 $f_1 \preccurlyeq f_2$ 意指

$$\mathrm{Dom}\, f_1 \subset \mathrm{Dom}\, f_2 \quad \text{及} \quad f_2(x) = f_1(x), \quad \text{对所有 } x \in \mathrm{Dom}\, f_1.$$

给定 $\mathcal{F}$ 的一个全序 (仍见 1.3 节) 子集 $\mathcal{E}$, 置

$$\mathrm{Dom}\, g := \bigcup_{f \in \mathcal{E}} \mathrm{Dom}\, f,$$

它显然是 X 的子空间, 此因 $\mathcal{E}$ 是全序的. 我们证明, 对任意 $x \in \mathrm{Dom}\, g$, 关系式

$$g(x) := f(x), \quad \text{对所有使得 } x \in \mathrm{Dom}\, f \text{ 的 } f \in \mathcal{E}$$

确切地定义一个线性泛函 $g : \mathrm{Dom}\, g \to \mathbb{R}$, 它对所有 $x \in \mathrm{Dom}\, g$ 满足 $g(x) \leqslant p(x)$.

现说明这一点. 设 $x \in \mathrm{Dom}\, g$ 使得 $x \in \mathrm{Dom}\, f_1$ 且 $x \in \mathrm{Dom}\, f_2$, 而 $f_1, f_2 \in \mathcal{E}$, 不妨假定 $f_1 \preccurlyeq f_2$ (子集 $\mathcal{E}$ 是全序的). 因此

$$g(x) = f_1(x) = f_2(x) \leqslant p(x).$$

如果 $x_1 \in \mathrm{Dom}\, f_1$ 而 $x_2 \in \mathrm{Dom}\, f_2$, 假如 $f_1 \preccurlyeq f_2$ 则 $(x_1 + x_2) \in \mathrm{Dom}\, f_2$ (如前述), 所以 $g(x_1 + x_2) = f_2(x_1 + x_2) = f_2(x_1) + f_2(x_2) = g(x_1) + g(x_2)$. 同样地, 对任意的 $\alpha \in \mathbb{R}, g(\alpha x_1) = f_1(\alpha x_1) = \alpha f_1(x_1) = \alpha g(x_1)$.

进而, g 显然是 $\mathcal{E}$ 的一个上界. 这是因为由构造知, $\mathrm{Dom}\, f \subset \mathrm{Dom}\, g$ 对所有 $f \in \mathcal{E}$ 成立. 由 Zorn 引理 (定理 1.3-1), 集合 $\mathcal{F}$ 具有极大元素 $\widetilde{\ell}$, 它定义在 X 的一个子空间 $\mathrm{Dom}\, \widetilde{\ell}$ 上.

然后可得

$$\mathrm{Dom}\, \widetilde{\ell} = X,$$

这意味着线性泛函 $\widetilde{\ell} \in \mathcal{F}$ 具有要求的所有性质. 如果不然, 则 $\widetilde{\ell} \subsetneqq X$, 如同在 (i) 那样同样地构造, 可得一线性泛函 $\widetilde{f} : \mathrm{Dom}\, \widetilde{f} \to \mathbb{R}$, 它满足

$$\mathrm{Dom}\, \widetilde{\ell} \subsetneqq \mathrm{Dom} \widetilde{f} \subset X, \quad \widetilde{f}(y) = \ell(y), \quad \text{对所有 } y \in Y,$$
$$\text{且 } \widetilde{f}(x) \leqslant p(x), \quad \text{对所有 } x \in \mathrm{Dom}\, \widetilde{f},$$

这与 $\widetilde{\ell}$ 的极大特征矛盾. 所以 $\mathrm{Dom}\, \widetilde{\ell} = X$. □

显然, 范数及半范数可作为次线性泛函 (如在定理 5.8-1 中所定义) 的例子. 但它们更一般, 因为数乘性质 $p(\alpha x) = |\alpha| p(x)$ 只要求对 $\alpha > 0$ 成立. 在 Hahn-Banach 定理的几何形式 (定理 5.10-1) 的证明中用到的 Minkowski 泛函提供了一个不是半范数的次线性泛函的例子.

习题

5.8-1 (复向量空间中的 Hahn-Banach 定理[23]) 设 X 是复向量空间, 而 $p : X \to \mathbb{R}$ 是 X 上的半范数 (因此较定理 5.8-1 中的假设, 受更多限制, 在那里 $p : X \to \mathbb{R}$ 只假定是次线性泛函). 又设 Y 是 X 的子空间, 而 $\ell : Y \to \mathbb{C}$ 是 Y 上的线性泛函, 它满足 $|\ell(y)| \leqslant p(y)$ 对所有 $y \in Y$ 成立.

证明在 X 上存在一个线性泛函 $\widetilde{\ell} : X \to \mathbb{C}$, 它满足 $\widetilde{\ell}(y) = \ell(y)$ 对所有 $y \in Y$ 且 $|\widetilde{\ell}(x)| \leqslant p(x)$ 对所有 $x \in X$ 成立.

提示: 对每一 $y \in Y$, 将 $\ell(y)$ 写为 $\ell(y) = \mathrm{Re}(\ell(y)) - i\mathrm{Re}(\ell(iy))$, 并注意到一个复向量空间当然也是一个实向量空间. 然后应用定理 5.8-1 于实线性泛函 $y \in Y \to \mathrm{Re}(\ell(y)) \in \mathbb{R}$ 及 $y \in Y \to \mathrm{Re}(\ell(iy)) \in \mathbb{R}$.

5.9 赋范向量空间的 Hahn-Banach 定理; 第一个推论

我们回顾一下 (3.5 节), 符号 X' 表示一个赋范向量空间 X 的对偶空间, 即由所有定义在 X 上的连续线性泛函 $\ell : X \to \mathbb{K}$ 构成的 Banach 空间, 而对任一 $\ell \in X'$, 其

[23] 这个定理归于:

H. F. Bohnenblust; A. Sobczyk [1938]: Extensions of functionals on complex linear spaces. Bulletin of the American Mathematical Society **44**, 91–93.

G. A. Soukhomlinoff [1938]: Über Fortsetzung von linearen Funktionalen in linearen komplexen Räumen und linearen Quaternionräumen. Mathematiceskii Sbornik **3**, 353–358.

范数由下式给出:

$$\|\ell\|_{X'} = \sup_{\begin{cases} x \in X \\ x \neq 0 \end{cases}} \frac{|\ell(x)|}{\|x\|_X}.$$

正像其在向量空间中的体现 (定理 5.8-1), 在赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1) 断言, 给定一个定义在 X 的子空间 Y 上的线性泛函 ℓ, 存在一个线性泛函 $\widetilde{\ell}$, 它是 ℓ 在全空间 X 上的延拓, 该延拓 $\widetilde{\ell}$ 与 ℓ 共具特定的性质. 这一性质在定理 5.8-1 中采取不等式的形式: $\ell(y) \leqslant p(y)$ 对所有 $y \in Y$ 及 $\widetilde{\ell}(x) \leqslant p(x)$ 对所有 $x \in X$ 成立; 现在则采用关系式 $\|\ell\|_{Y'} = \|\widetilde{\ell}\|_{X'}$ 的形式.

要强调的是, 本节在实情况成立的所有定理可逐字逐句地转移至复情况.

下面, 像 ℓ 或 ℓ_x 这类符号被用来表示对偶空间 X' 的特定元素, 而符号 x' 通常被用来表示 X' 的一般元素.

定理 5.9-1 (赋范向量空间中的 Hahn-Banach 定理) *设 X 是赋范向量空间, Y 是 X 的子空间, 另设 $\ell : Y \to \mathbb{K}$ 是连续线性泛函.*

则存在一个连续线性泛函 $\widetilde{\ell} : X \to \mathbb{K}$, 满足

$$\widetilde{\ell}(y) = \ell(y), \quad \text{对所有 } y \in Y \text{ 且 } \|\widetilde{\ell}\|_{X'} = \|\ell\|_{Y'}.$$

证明 设 X 是实赋范向量空间. 由 $p(x) := \|\ell\|\|x\|$ 对所有 $x \in X$, 其中 $\|\ell\| := \|\ell\|_{Y'}$, 定义的函数 $p : X \to \mathbb{R}$ 是一个范数 (除非 $\ell = 0$), 因此是一个 X 上的次线性泛函; 此外

$$\ell(y) \leqslant \|\ell\|\|y\| = p(y), \quad \text{对所有 } y \in Y.$$

由实向量空间的 Hahn-Banach 定理 (定理 5.8-1), 存在一个线性泛函 $\widetilde{\ell} : X \to \mathbb{R}$ 满足

$$\begin{aligned} &\widetilde{\ell}(y) = \ell(y), \quad \text{对所有 } y \in Y, \\ &\widetilde{\ell}(x) \leqslant p(x) = \|\ell\|\|x\| \quad \text{且} \quad -\widetilde{\ell}(x) = \widetilde{\ell}(-x) \leqslant p(-x) = \|\ell\|\|x\|, \quad \text{对所有 } x \in X. \end{aligned}$$

因此 $|\widetilde{\ell}(x)| \leqslant \|\ell\|\|x\|$ 对所有 $x \in X$ 成立. 故得, 线性泛函 $\widetilde{\ell}$ 是连续的, 而且

$$\|\ell\| = \sup_{\begin{cases} y \in Y \\ y \neq 0 \end{cases}} \frac{|\ell(y)|}{\|y\|} \leqslant \sup_{\begin{cases} x \in X \\ x \neq 0 \end{cases}} \frac{|\widetilde{\ell}(x)|}{\|x\|} = \|\widetilde{\ell}\|_{X'} \leqslant \|\ell\|.$$

所以 $\|\widetilde{\ell}\|_{X'} = \|\ell\|_{Y'}$.

复情况的证明留给读者 (习题 5.9-1). □

注 在 Hilbert 空间中的 Hahn-Banach 定理, 可以利用直接和定理以及 F. Riesz 表示定理以简单得多的方式予以证明, 而且还能给出延拓的唯一性 (定理 4.7-1). 在这种情况下, 无须再借助于选择公理.

需要强调的是, 一般而言这种保范延拓 $\widetilde{\ell}$ 不一定是唯一的. 例如, 置 $X = \mathcal{P}[0,1]$ 装备以 sup 范数 $\|\cdot\|, Y = \mathcal{P}_3[0,1]$, 而线性泛函 $\ell : Y \to \mathbb{R}$ 由下式定义:

$$\ell(p) = \frac{1}{6}\left(p(0) + 4p\left(\frac{1}{2}\right) + p(1)\right), \quad \text{对所有 } p \in \mathcal{P}_3[0,1].$$

则 ℓ 是连续的, $\|\ell\| = \sup_{\left\{\begin{smallmatrix} p \in \mathcal{P}_3[0,1] \\ p \neq 0 \end{smallmatrix}\right.} \dfrac{|\ell(p)|}{\|p\|} = 1.$

然而可以立即验证, 由下式定义的两个不同的线性形式 $\widetilde{\ell}_1 : X \to \mathbb{R}$ 及 $\widetilde{\ell}_2 : X \to \mathbb{R}$

$$\widetilde{\ell}_1(f) = \frac{1}{6}\left(f(0) + 4f\left(\frac{1}{2}\right) + f(1)\right) \text{ 及}$$

$$\widetilde{\ell}_2(f) = \int_0^1 f(t)dt, \quad \text{对所有 } f \in \mathcal{P}[0,1]$$

都满足

$$\widetilde{\ell}_\alpha(p) = \ell(p), \quad \text{对所有 } p \in \mathcal{P}_3[0,1] \text{ 及}$$

$$\|\widetilde{\ell}_\alpha\| = \sup_{\left\{\begin{smallmatrix} f \in \mathcal{C}[0,1] \\ f \neq 0 \end{smallmatrix}\right.} \frac{|\widetilde{\ell}_\alpha(f)|}{\|f\|} = 1, \quad \text{对 } \alpha = 1, 2.$$

然而, 根据下面的结果, 有一大类的赋范向量空间, 其中连续线性泛函的保范延拓是唯一的. 回顾一下, 一个实或复赋范向量空间是严格凸的 (2.17 节), 如果

$$x \neq y \text{ 且 } \|x\| = \|y\| = 1 \quad \text{意味着} \quad \left\|\frac{x+y}{2}\right\| < 1.$$

定理 5.9-2 (Taylor-Foguel 定理[24]) 设 X 是赋范向量空间. 则所有定义在 X 子空间上的连续线性泛函有到 X 上的唯一保范延拓, 其充分必要条件是 X 的对偶空间 X' 是严格凸的.

证明　我们给出实情况的证明, 复情况留作习题 (习题 5.9-3).

(i) 假定存在 X 的一个子空间 Y 及连续线性泛函 $\ell \in Y'$ 和 $\ell_1, \ell_2 \in X'$ 使得

$$\ell_1(y) = \ell_2(y) = \ell(y), \quad \text{对所有 } y \in Y,$$

$$\|\ell_1\|_{X'} = \|\ell_2\|_{X'} = \|\ell\|_{Y'} = 1 \quad \text{且 } \ell_1 \neq \ell_2$$

[24] A. E. Taylor [1939]: The extension of linear functionals. Duke Mathematical Journal **5**, 538–547.

S. R. Foguel [1958]: On a theorem of A.E. Taylor. Proceedings of the American Mathematical Society **9**, 325.

这里给出的简单证明取自:

P. R. Beesack; E. Hughes; M. Ortel [1979]: Rotund complex linear spaces. Proceedings of the American Mathematical Society **75**, 42–44.

Taylor-Foguel 定理也可以作为更一般的 Phelps 定理的推论导出, 见:

R. Phelps [1960]: Uniqueness of Hahn-Banach extensions and unique best approximation. Transactions of the American Mathematical Society **95**, 238–255.

(显然, 假设 $\|\ell\|_{Y'}=1$ 并不失一般性). 那么, 因为 $\frac{1}{2}(\ell_1+\ell_2)(y)=\ell(y)$ 对所有 $y\in Y$ 成立, 就有

$$1=\|\ell\|_{Y'}=\left\|\frac{\ell_1+\ell_2}{2}\right\|_{Y'}\leqslant\left\|\frac{\ell_1+\ell_2}{2}\right\|_{X'}\leqslant 1,$$

这说明 X' 不是严格凸的.

(ii) 假设定义在 X 的子空间上的连续线性泛函到 X 上的所有保范延拓都是唯一的.

设 $\ell_1,\ell_2\in X'$ 使得

$$\|\ell_1\|_{X'}=\|\ell_2\|_{X'}=1 \quad 且 \quad \ell_1\neq\ell_2.$$

则

$$Y:=\{y\in X;\ell_1(y)=\ell_2(y)\}$$

是 X 的真子空间, 且由 $\ell(y):=\ell_1(y)=\ell_2(y)$ 对所有 $y\in Y$ 定义的连续线性泛函 $\ell\in Y'$ 就满足

$$\|\ell\|_{Y'}<1.$$

为说明这一点, 注意 $\|\ell\|_{Y'}\leqslant\|\ell_1\|_{X'}=1$, 而且如果 $\|\ell\|_{Y'}=1$, 则 ℓ_1 与 ℓ_2 将相等, 得到矛盾.

由假设, 存在唯一的 $\widetilde{\ell}\in X'$ 使得

$$\widetilde{\ell}(y)=\ell(y)=\ell_1(y)=\ell_2(y), \quad 对所有\ y\in Y\ 且\ \|\widetilde{\ell}\|_{X'}=\|\ell\|_{Y'}<1.$$

设 $x_0\in X-Y$. 那么因为 $\ell_1(x_0)\neq\ell_2(x_0)$, 存在 $\lambda=\lambda(x_0)\in\mathbb{R}$ 和 $\mu=\mu(x_0)\in\mathbb{R}$ 使得

$$\lambda\ell_1(x_0)+\mu\ell_2(x_0)=\widetilde{\ell}(x_0) \quad 及\ \lambda+\mu=1.$$

这就得到 $\widetilde{\ell}(x_0)=\lambda\ell_1(x_0)+(1-\lambda)\ell_2(x_0)$, 而且

$$\widetilde{\ell}=\lambda\ell_1+(1-\lambda)\ell_2,$$

这是因为每个 $x\in X$ 均可写为 $x=y+\alpha x_0$, 对某个 $y\in Y$ 及 $\alpha\in\mathbb{R}$.

我们现说明 $\lambda\in]0,1[$. 如果 $\lambda\geqslant 1$, 则 $\ell_1=\frac{1}{\lambda}\widetilde{\ell}+\frac{\lambda-1}{\lambda}\ell_2$ 意味着 $\|\ell_1\|_{X'}<1$, 矛盾; 而如果 $\lambda\leqslant 0$, 则 $\ell_2=\frac{1}{1-\lambda}\widetilde{\ell}-\frac{\lambda}{1-\lambda}\ell_1$ 意味着 $\|\ell_2\|_{X'}<1$, 也得矛盾. 所以 $\lambda\in]0,1[$.

如果 $0<\lambda<\frac{1}{2}$, 则 $\frac{\ell_1+\ell_2}{2}=\frac{1}{2(1-\lambda)}\widetilde{\ell}+\frac{1-2\lambda}{2(1-\lambda)}\ell_1$ 意味着 $\left\|\frac{\ell_1+\ell_2}{2}\right\|_{X'}<1$; 如果 $\frac{1}{2}<\lambda<1$, 则 $\frac{\ell_1+\ell_2}{2}=\frac{1}{2\lambda}\widetilde{\ell}+\left(1-\frac{1}{2\lambda}\right)\ell_2$ 意味着 $\left\|\frac{\ell_1+\ell_2}{2}\right\|_{X'}<1$; 最后, 如果 $\lambda=\frac{1}{2}$, 则 $\left\|\frac{\ell_1+\ell_2}{2}\right\|_{X'}=\|\widetilde{\ell}\|_{X'}<1$. 所以 X' 是严格凸的. □

注 如果 X 是 Hilbert 空间, 利用 F. Riesz 表示定理, 已经证明了 Taylor-Foguel 定理的“充分性”部分 (定理 4.7-1).

借助于赋范空间的 Hahn-Banach 定理, 我们现在能够 (正面地) 回答关于定义在任意赋范向量空间 X 上非零连续线性泛函存在性的问题.

定理 5.9-3 设 $X \neq \{0\}$ 是赋范向量空间. 给定任意的非零向量 $x \in X$, 存在 $\ell_x \in X'$ 使得

$$\ell_x(x) = \|x\| \quad 且 \quad \|\ell_x\|_{X'} = 1.$$

也就是说, 对偶空间 X' 包含非零元素.

证明 设

$$Y := \{\alpha x \in X; \alpha \in \mathbb{K}\} = \mathrm{Span}(x) \quad 及 \quad \ell(\alpha x) := \alpha\|x\|, \quad 对任意\ x \in \mathbb{K},$$

如此定义的泛函 $\ell : Y \to \mathbb{R}$ 是 X 的子空间 Y 上的连续线性泛函, 满足

$$\|\ell\|_{Y'} = \sup_{\alpha \neq 0} \frac{|\ell(\alpha x)|}{\|\alpha x\|} = 1.$$

则由定理 5.9-1 知存在连续线性泛函 $\ell_x \in X'$ 使得

$$\ell_x(\alpha x) = \ell(\alpha x) = \alpha\|x\|, \quad 对所有\ \alpha \in \mathbb{R}\ 且\ \|\ell_x\| = \|\ell\| = 1.$$

因为 $\ell(x) = \|x\|$, 所以泛函 ℓ_x 具有所要求的性质. □

由定理 5.9-3, 给定任意的 $x_0 \in X, \|x_0\| = 1$, 存在 $x' \in X'$ 使得 $x'(x_0) = 1$ 且 $\|x'\| = 1$. 所以, 对这样的 $x' \in X'$, 有

$$\|x'\| = \sup_{\|x\|=1} |x'(x)| = |x'(x_0)|,$$

即定义 x' 的范数的上确界是可以达到的. 实际上, 有“许多”这种连续线性泛函, 而这正是线性泛函分析的另一基本定理的内容.

定理 5.9-4 (Bishop-Phelps 定理[25]) 设 X 是实 Banach 空间, 又设

$$Y' := \{x' \in X';\ 存在\ x_0\ 使得\ \|x_0\| = 1\ 且\ \sup_{\|x\|=1} |x'(x)| = |x'(x_0)|\}.$$

则 Y' 在 X' 中稠密.

[25] E. Bishop; R. R. Phelps [1961]: A proof that every Banach space is subreflexive. Bulletin of the American Mathematical Society **67**, 97–98.

这个结果可以扩展到具有“Radon-Nikodym 性质”的复 Banach 空间, 可参阅:

J. Bourgain [1977]: On dentability and the Bishop-Phelps property. Israel Journal of Mathematics **28**, 268–271.

在这个方向上, 另一个深刻的定理断言, 如果一个实 Banch 空间 X 满足 $Y' = X'$, 其中 Y' 如定理 5.9-4 中定义, 则 X 是自反的 (这个概念将在 5.14 节中定义); 这个结果属于:

R. C. James [1972]: Reflexivity and the sup of linear functionals. Israel Journal of Mathematics **13**, 289–301.

赋范向量空间 X 中的 Hahn-Banach 定理 (通过其推论, 定理 5.9-3) 的另一重要结论是, 任一元素 $x \in X$ 的范数 $\|x\|$ 可由与以下公式互反的一个公式给出:

$$\|x'\| = \sup_{\left\{\substack{x \in X \\ x \neq 0}\right.} \frac{|x'(x)|}{\|x\|}, \quad \text{对任意 } x' \in X'$$

("互反" 一词的含义是 X 与 X' 的角色可以互换). 这个公式还转而回答了这样一个问题: 对偶空间 X' 是否包含 "足够" 的元素 x' 以保证 $x'(x) = 0$ 对所有 $x' \in X'$ 成立意味着 $x = 0$? 注意, 非零元素 $x' \in X'$ 的存在 (定理 5.9-3) 确保了出现在下一定理中的上确界是适定的.

定理 5.9-5 设 $X \neq \{0\}$ 是一个赋范向量空间. 则任一向量 $x \in X$ 的范数由下式给出:

$$\|x\| = \sup_{\left\{\substack{x' \in X' \\ x' \neq 0}\right.} \frac{|x'(x)|}{\|x'\|} = \sup_{\left\{\substack{x' \in X' \\ \|x'\| = 1}\right.} |x'(x)|.$$

所以, 如果 $x \in X$ 使得 $x'(x) = 0$ 对所有 $x' \in X'$ 成立, 则 $x = 0$.

证明 给定任意 $x \in X$, 设 $\ell_x \in X'$ 是定理 5.9-3 中所确定的元素, 即 $\ell_x(x) = \|x\|$ 和 $\|\ell_x\| = 1$. 这就有

$$\|x\| = \ell_x(x) \leqslant \sup_{\left\{\substack{x' \in X' \\ \|x'\| = 1}\right.} |x'(x)| \leqslant \|x\|. \qquad \square$$

注 在 Hilbert 空间 $(X, (\cdot, \cdot))$ 中, 利用 F. Riesz 等距 (定理 4.6-1), X 的任一元素均可等同于其对偶空间 X' 的一个元素. 出于这种考虑, 定理 5.9-5 中确立的关系式 $\|x\| = \sup_{\left\{\substack{x' \in X' \\ x' \neq 0}\right.} \dfrac{|x'(x)|}{\|x'\|}$ 在这种情况下就等价于关系式 $\|x\| = \sup_{y \neq 0} \dfrac{|(x, y)|}{\|y\|}$ (实际上, 这对任何完备或不完备的内积空间都成立; 见定理 4.1-1).

赋范向量空间中 Hahn-Banach 定理的另一个重要结论是将在定理 5.9-8 中给出的、关于赋范向量空间可分性的一个简单的充分条件. 为此目的, 我们首先要证明定理 5.9-3 (可视为下述定理在 $Y = \{0\}$ 的特殊情况) 本身的一个有趣推广. 特别地, 它证明了如果 Y 是赋范向量空间 X 的一个闭且严格的子空间, 一定存在一个 X 上的非零连续泛函, 在 Y 上为零.

定理 5.9-6 设 X 是赋范向量空间, 而 Y 是 X 的一个闭子空间且 $Y \subsetneq X$. 给定任意元素 $x \in X - Y$, 一定存在 $\ell_x \in X'$ 使得

$$\ell_x(y) = 0, \quad \text{对所有 } y \in Y, \quad \ell_x(x) = \inf_{y \in Y} \|x - y\| > 0 \text{ 且 } \|\ell_x\| = 1.$$

证明 由以下诸式定义 X 的子空间 Z 及函数 $\ell : Z \to \mathbb{K}$:

$$Z := \{(\alpha x + y) \in X; \alpha \in \mathbb{K}, y \in Y\},$$
$$\ell(\alpha x + y) := \alpha\delta, \quad \text{对所有 } \alpha \in \mathbb{K} \text{ 及 } y \in Y,$$

其中 $\delta := \inf_{y\in Y} \|x - y\| > 0$. 显然, ℓ 是 Z 上的线性泛函且满足 $\ell(x) = \delta$ 及 $\ell(y) = 0$ 对所有 $y \in Y$ 成立. 进而, 我们断言 ℓ 是连续的而且 $\|\ell\|_{Z'} = 1$. 为说明这一点, 我们首先注意, 由 δ 的定义, 任一元素 $(\alpha x + y) \in Z, \alpha \neq 0$ 满足

$$\|\alpha x + y\| = |\alpha|\|x + \alpha^{-1}y\| \geqslant |\alpha|\delta,$$

此因 $\alpha^{-1}y \in Y$, 从而 $|\ell(\alpha x + y)| = |\alpha|\delta \leqslant \|\alpha x + y\|$. 又注意到这个不等式对 $\alpha = 0$ 也成立, 所以有

$$\|\ell\|_{Z'} = \sup_{\left\{\substack{\alpha \in \mathbb{K}, y \in Y \\ \alpha x + y \neq 0}\right.} \frac{|\ell(\alpha x + y)|}{\|\alpha x + y\|} \leqslant 1.$$

其次, 注意到仍由 δ 的定义, 对任意 $\varepsilon > 0$, 存在 $y_\varepsilon \in Y$ 使得 $\delta \leqslant \|x - y_\varepsilon\| \leqslant \delta + \varepsilon$. 这样, 由于 $(x - y_\varepsilon) \in Z$ 且 $\ell(x - y_\varepsilon) = \ell(x) = \delta$, 我们有

$$1 \geqslant \|\ell\|_{Z'} = \sup_{\left\{\substack{z \in Z \\ z \neq 0}\right.} \frac{|\ell(z)|}{\|z\|} \geqslant \frac{|\ell(x - y_\varepsilon)|}{\|x - y_\varepsilon\|} \geqslant \frac{\delta}{\delta + \varepsilon}.$$

因 $\varepsilon > 0$ 是任意的, 所以 $\|\ell\|_{Z'} = 1$.

然后由赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1), 存在一个连续线性泛函 $\ell_x \in X'$ 满足

$$\ell_x(z) = \ell(z), \quad \text{对所有 } z \in Z \text{ 且 } \|\ell_x\| = \|\ell\| = 1.$$

特别地, $\ell_x(x) = \ell(x) = \delta$ 且 $\ell_x(y) = \ell(y) = 0$ 对所有 $y \in Y$ 成立. □

注 粗略地扫视一下定理的证明可见, 即使子空间 Y 不是闭的, 定理 5.9-6 对任意满足 $\inf_{y\in Y} \|x - y\| > 0$ 的 $x \in (X - Y)$ 仍然成立.

定理 5.9-6 正是内积空间 X 中的投影定理 (定理 4.3-1) 在任意赋范向量空间中的推广. 事实上, 如果 Z 是 X 的完备子空间而 $x \in X - Z$, 投影定理断言, 存在一个元素 $Px \in Z$ 使得 $\|x - Px\| = \inf_{z\in Z} \|x - z\|$, 它进而还具有以下特性: $(x - Px, z) = 0$ 对所有 $z \in Z$ 成立. 而根据 F. Riesz 表示定理, 元素

$$\ell_x := \frac{x - Px}{\|x - Px\|} \in X$$

等同于一个元素 $\ell_x \in X'$, 它严格地满足在定理 5.9-6 中展现的那些性质, 即

$$\begin{aligned}
&\|\ell_x\| = 1,\\
&\ell_x(z) = \frac{1}{\|x - Px\|}(x - Px, z) = 0, \quad \text{对所有 } z \in Z,\\
&\ell_x(x) = \frac{1}{\|x - Px\|}(x - Px, x)\\
&\qquad = \frac{1}{\|x - Px\|}(x - Px, x - Px + Px) = \|x - Px\|.
\end{aligned}$$

定理 5.9-6 的第一个推论是下述判定一个子空间在赋范向量空间中是否稠密的准则.

定理 5.9-7 设 Y 是赋范向量空间 X 的子空间, 则 $\overline{Y} = X$ 的充分必要条件是

$$\{x' \in X'; x'(y) = 0, \text{ 对所有 } y \in Y\} = \{0\}.$$

证明 若 $\overline{Y} = X, x' \in X'$ 使得 $x'(y) = 0$ 对所有 $y \in Y$ 成立, 则 $x'(y) = 0$ 对所有 $y \in Y$, 进而 $x'(y) = 0$ 对所有 $y \in \overline{Y} = X$ 成立 (此因 $x' : X \to \mathbb{K}$ 是连续的). 所以 $x' = 0$.

若 $\overline{Y} \subsetneqq X$, 则由定理 5.9-6, 存在非零连续线性形式 $x' \in X'$ 使得 $x'(y) = 0$ 对所有 $y \in \overline{Y}$, 当然对所有 $y \in Y$ 成立. 所以在这种情况, $\{x' \in X'; x'(y) = 0, \text{对所有 } y \in Y\} \supsetneqq \{0\}$. □

注 如果 X 是 Hilbert 空间, 相应的 "充分性" 部分就是 F. Riesz 表示定理再加上投影定理 (定理 4.3-2) 的一个直接推论.

可分赋范向量空间的对偶不一定是可分的, 这可以空间 ℓ^1 及 $L^1(\Omega)$ 作为例证. 这两个空间的对偶可分别等同于空间 ℓ^∞ 和 $L^\infty(\Omega)$, 所以不是可分的 (定理 2.4-2, 2.5-4, 3.5-1 及 3.5-3). 但由定理 5.9-6 可得, 其逆却成立:

定理 5.9-8 (可分性的充分条件) 如果赋范向量空间 X 的对偶空间 X' 是可分的, 则 X 也是可分的.

证明 (i) 设 $S' := \{x' \in X'; \|x'\| = 1\}$. 则存在元素 $x'_n \in S', n \geqslant 1$, 使得

$$\overline{\cup_{n=1}^{\infty}\{x'_n\}} = S'.$$

设给定 $\varepsilon > 0$ 及 $x' \in S'$. 由假设 X' 是可分的, 故存在元素 $\widetilde{x}'_n \in X', n \geqslant 1$, 使得 $\overline{\cup_{n=1}^{\infty}\{\widetilde{x}'_n\}} = X'$. 给定任意的 $0 < \varepsilon < 2$ 及任意的 $x' \in S'$, 存在整数 $n \geqslant 1$ 使得 $\|\widetilde{x}'_n - x'\| \leqslant \frac{\varepsilon}{2}$. 所以

$$|\|\widetilde{x}'_n\| - 1| = |\|\widetilde{x}'_n\| - \|x'\|| \leqslant \|\widetilde{x}'_n - x'\| \leqslant \frac{\varepsilon}{2},$$

故 $0 < 1 - \dfrac{\varepsilon}{2} \leqslant \|\widetilde{x}'_n\|$. 设 $x'_n := \dfrac{\widetilde{x}'_n}{\|\widetilde{x}'_n\|}$, 则

$$\begin{aligned}\|x' - x'_n\| &\leqslant \|x' - \widetilde{x}'_n\| + \|\widetilde{x}'_n - x'_n\| \\ &= \|x' - \widetilde{x}'_n\| + (\|\widetilde{x}'_n\| - 1)\|x'_n\| \leqslant \varepsilon.\end{aligned}$$

这就证明了 $\overline{\cup_{n=1}^{\infty}\{x'_n\}} = S'$.

(ii) 设 $x'_n \in S' \subset X', n \geqslant 1$, 是出现在 (i) 中的泛函. 因为 $\|x'_n\| = \sup_{\|x\|=1} \|x'_n(x)\|$, 故存在 $x_n \in X$ 使得

$$\|x_n\| = 1 \quad 及 \quad \frac{1}{2} \leqslant |x'_n(x_n)|, \quad 对每个\ n \geqslant 1.$$

下面我们将证明

$$\overline{Y} = X, \quad 其中\ \overline{Y} := \mathrm{Span}(x_n)_{n=1}^{\infty},$$

这就证明了 X 是可分的, 因为这个关系式意味着向量 $x_n, n \geqslant 1$, 具有有理数系数的有限线性组合形成 X 的一个稠密子集.

为此目的, 我们用反证法. 如果 $Y \subsetneqq X$, 设 $x \in X - Y$. 由定理 5.9-6, 存在 $\ell_x \in X'$ 使得

$$\ell_x(y) = 0, \quad 对所有\ y \in Y\ 且\ \|\ell_x\| = 1$$

($\ell_x(x) = \inf_{y \in Y} \|x - y\|$ 这一性质在此并不需要). 因此 $\ell_x \in S'$ 而且

$$\begin{aligned}\frac{1}{2} < |x'_n(x_n)| &= |x'_n(x_n) - \ell_x(x_n)| \\ &\leqslant \|x'_n - \ell_x\|, \quad 对所有\ n \geqslant 1,\end{aligned}$$

这与在 (i) 中确立的 $\cup_{n=1}^{\infty}\{x'_n\}$ 在 S' 中的稠密性矛盾. □

注 当然, 类似于 (i) 的性质在任何可分赋范向量空间中均成立 (即不管它是否是对偶空间).

习题

5.9-1 对 X 是复赋范向量空间的情形, 证明定理 5.9-1.

提示: 应用复向量空间中的 Hahn–Banach 定理 (习题 5.8-1).

5.9-2 设 X 是赋范向量空间, Y 是 X 的严格子空间, 又设 $\ell: Y \to \mathbb{K}$ 是连续线性泛函. 证明存在连续线性泛函 $\widetilde{\ell}: X \to \mathbb{K}$ 满足

$$\widetilde{\ell}(y) = \ell(y), \quad 对所有\ y \in Y\ 且\ \|\widetilde{\ell}\|_{X'} > \|\ell\|_{Y'}.$$

5.9-3 在复情况下, 证明 Taylor–Foguel 定理 (定理 5.9-2).

提示: 首先, 证明一个复赋范向量空间 X 是严格凸的充分必要条件是, 对每对 $x, y \in X, x \neq y$ 且 $\|x\| = \|y\| = 1$, 存在 $\lambda \in \mathbb{C}$ 使得 $\|\lambda x + (1 - \lambda)y\| < 1$ (对 "充分性" 部分, 用定理 5.9-2). 然后修改书中给出的证明使之适合于复情况.

5.10 Hahn-Banach 定理的几何形式: 凸集的分离

给定一个实赋范向量空间 X, 一个非零线性连续泛函 $\ell: X \to \mathbb{R}$ 以及一个数 $\gamma \in \mathbb{R}$, 集合

$$\{x \in X; \ell(x) = \gamma\}$$

称作一个闭仿射超平面, 而集合 $\{x \in X; \ell(x) \geqslant \gamma\}$ 及 $\{x \in X; \ell(r) > \gamma\}$ 则分别称为闭及开半空间 (关于超平面及其与线性泛函的关系的补充材料在习题 5.10-1 及 5.10-2 中给出).

X 的两个集合 A 与 B 称为被超平面分离的, 如果存在一个非零 $\ell \in X'$ 与 $\gamma \in \mathbb{R}$ 使得

$$A \subset \{x \in X; \ell(x) \leqslant \gamma\} \quad \text{而} \quad B \subset \{y \in X; \gamma \leqslant \ell(y)\},$$

即它们分别包含在两个闭半空间内, 其交集是闭仿射超平面 $\{y \in X; \ell(y) = \gamma\}$ (图 5.10-1).

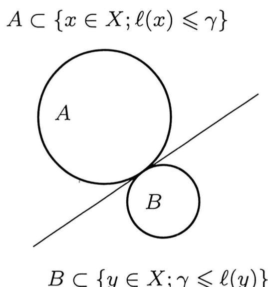


图 5.10-1　被超平面 (在此图中为一直线) 分离的 $\mathbb{R}^2$ 的两子集 A 与 B

下面的定理紧密地依存于向量空间的 Hahn-Banach 定理 (定理 5.8-1), 它给出了实赋范向量空间的两个集合被一个超平面分离的充分条件. 对于扩展到复空间的情况, 见习题 5.10-3(1).

定理 5.10-1 (Hahn-Banach 定理的第一几何形式: 凸集的分离) 设 A 与 B 为实赋范向量空间 X 的两个非空子集, 其满足以下性质:

$$A \text{ 是开凸集}; \quad B \text{ 是凸集}; \quad A \cap B = \varnothing.$$

那么存在非零的 $\ell \in X'$ 及 $\gamma \in \mathbb{R}$ 使得 (图 5.10-1)

$$\ell(x) \leqslant \gamma \leqslant \ell(y), \quad \text{对所有 } x \in A, y \in B.$$

证明　(i) 设 C 是 X 的包含原点的非空凸开子集. 那么对每一个 $x \in X$, 由

$$p(x) := \inf\left\{\beta > 0; \frac{x}{\beta} \in C\right\}$$

定义的函数 $p: X \to [0, \infty)$ 具有以下 4 条性质:

存在一个常数 M 使得

$$\begin{aligned}
&0 \leqslant p(x) \leqslant M\|x\|, \quad \text{对所有 } x \in X, \\
&C = \{x \in X; p(x) < 1\}, \\
&p(\alpha x) = \alpha p(x), \quad \text{对所有 } \alpha \geqslant 0, x \in X, \\
&p(x+y) \leqslant p(x) + p(y), \quad \text{对所有 } x, y \in X.
\end{aligned}$$

因为 C 是开的且包含原点, 故存在 $r > 0$ 使 $B(0; r) \subset C$. 给定任意 $x \in X$ 及任意的 $\beta > \frac{\|x\|}{r}$, 点 $\frac{x}{\beta}$ 属于 $B(0; r)$ (因 $\left\|\frac{x}{\beta}\right\| < r$), 故 $\frac{x}{\beta} \in C$. 所以

$$p(x) = \inf\left\{\beta > 0; \frac{x}{\beta} \in C\right\} \leqslant \inf\left\{\beta > \frac{\|x\|}{r}\right\} = \frac{\|x\|}{r}.$$

因此有 $0 \leqslant p(x) \leqslant M\|x\|$ 对所有 $x \in X$ 成立, 其中 $M := \frac{1}{r}$.

给定任意 $x \in C$. 因 C 是开的, 存在 $\delta > 0$ 使 $(1+\delta)x \in C$. 这就有

$$p(x) = \inf\left\{\beta > 0; \frac{x}{\beta} \in C\right\} \leqslant \frac{1}{1+\delta} < 1.$$

反之, 设 $x \in X$ 且 $p(x) < 1$. 这意味着存在 $0 < \beta < 1$ 使得 $\frac{x}{\beta} \in C$. 因为 C 是凸的且包含原点, 故 $x = \left(\beta\frac{x}{\beta} + (1-\beta)0\right) \in C$. 所以 $C = \{x \in X; p(x) < 1\}$.

给定任意 $\alpha > 0$ 及任意的 $x \in X$,

$$\begin{aligned}
p(\alpha x) &= \inf\left\{\widetilde{\beta} > 0; \frac{\alpha x}{\widetilde{\beta}} \in C\right\} \\
&= \inf\left\{\alpha\beta > 0; \frac{\alpha x}{\alpha\beta} = \frac{x}{\beta} \in C\right\} = \alpha p(x).
\end{aligned}$$

此外, $p(0) = 0$. 所以 $p(\alpha x) = \alpha p(x)$ 对所有 $\alpha \geqslant 0$ 及所有 $x \in X$ 成立.

最后, 设两点 $x, y \in X$ 及 $\varepsilon > 0$ 是给定的. 由上述性质有

$$p\left(\frac{x}{p(x)+\varepsilon}\right) = \frac{p(x)}{p(x)+\varepsilon} < 1.$$

因此 $\left(\frac{x}{p(x)+\varepsilon}\right) \in C$; 同理 $\left(\frac{y}{p(y)+\varepsilon}\right) \in C$. 这样, C 的凸性意味着

$$\left(\frac{\mu}{p(x)+\varepsilon}x + \frac{1-\mu}{p(y)+\varepsilon}y\right) \in C$$

对所有 $0 < \mu < 1$ 成立. 特别地选取 $\mu := \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$, 就有

$$\frac{1}{p(x)+p(y)+2\varepsilon}(x+y) \in C.$$

这就得到

$$p(x+y) < p(x) + p(y) + 2\varepsilon.$$

因为 $\varepsilon > 0$ 是任意的, 所以 $p(x+y) \leqslant p(x) + p(y)$.

这就证明了函数 p 具有所列出的所有性质. 特别地, 后两条性质说明 p 是一个次线性泛函 (5.8 节).

(ii) 设 C 是 X 的一个非空凸开子集, 而 $y_0 \notin C$. 则存在 $\ell \in X'$ 使得

$$\ell(x) < \ell(y_0), \quad \text{对所有 } x \in C$$

(注意, (ii) 实际上是定理 5.10-1 的一个特殊情况, 其中 $A := \{y_0\}, B := C$).

先假定 $O \in C$. 设 $p : X \to [0, \infty[$ 为由 (i) 中所定义的函数, $Y := \text{Span}\{y_0\}$ (因 $O \in C$, 而 $y_0 \notin C$, 故 y_0 是非零向量), 而 $\ell_0 : Y \to \mathbb{R}$ 为如下定义的线性泛函: 对每个 $\alpha \in \mathbb{R}, \ell_0(\alpha y_0) = \alpha$. 这就有

$$\ell_0(y) \leqslant p(y), \quad \text{对所有 } y \in Y,$$

此因

$$\ell_0(\alpha y_0) = \alpha \leqslant \alpha p(y_0) = p(\alpha y_0), \quad \text{对所有 } \alpha \geqslant 0$$

(由 (i), 当 $y_0 \notin C$ 时 $p(y_0) \geqslant 1$), 及

$$\ell(\alpha y_0) = \alpha \leqslant 0 \leqslant p(\alpha_0 y), \quad \text{对所有 } \alpha \leqslant 0$$

(由 (i), $p(y) \geqslant 0$ 对所有 $y \in X$ 成立).

因为由 (i), p 是一个次线性泛函, 实向量空间中的 Hahn-Banach 定理 (定理 5.8-1) 说明, 存在一个非零线性泛函 $\ell : X \to \mathbb{R}$ 使得

$$\ell(y_0) = \ell_0(y_0) = 1 \quad \text{且} \quad \ell(x) \leqslant p(x), \quad \text{对所有 } x \in X.$$

进而, 在 (i) 中确立的对所有 $x \in X$ 成立不等式 $p(x) \leqslant M\|x\|$ 意味着

$$\ell(x) \leqslant p(x) \leqslant M\|x\| \quad \text{及} \quad -\ell(x) \leqslant p(-x) \leqslant M\|x\|, \quad \text{对所有 } x \in X.$$

所以线性泛函 ℓ 是连续的, 即 $\ell \in X'$. 此外,

$$\ell(x) \leqslant p(x) < 1 = \ell_0(y_0) = \ell(y_0), \quad \text{对所有 } x \in C$$

(由 (i), $p(x) < 1$ 对所有 $x \in C$ 成立). 所以如果 $O \in C$, 上述断言就得证.

现在假定 $O \notin C$. 任意选取点 $x_0 \in C$, 令

$$\widetilde{C} := \{(x - x_0) \in X; x \in C\} \quad \text{及} \quad \widetilde{y}_0 := y_0 - x_0.$$

因为 $O \in \widetilde{C}$ 且 $\widetilde{y}_0 \notin \widetilde{C}$, 前面的推导说明, 存在 $\ell \in X'$ 使得

$$\ell(\tilde{x}) < \ell(\widetilde{y}_0), \quad 对所有\ \widetilde{x} \in \widetilde{C},$$

由于 ℓ 是线性的, 因此 $\ell(x) < \ell(y_0)$ 对所有 $x \in C$ 成立.

(iii) 最后, 设 A 是 X 的非空凸开子集. B 是 X 的非空凸子集使得 $A \cap B = \varnothing$ (如定理中所述).

定义集合

$$C := \bigcup_{y \in B} \{(x-y) \in X; x \in A\},$$

作为开集的并, 它是开集. 实际上它也是凸的: 设 $z_i = (x_i - y_i) \in C$, 其中 $x_i \in A, y_i \in B, i = 1, 2$. 因为

$$(\mu x_1 + (1-\mu)x_2) \in A, \quad (\mu y_1 + (1-\mu)y_2) \in B, \quad 对所有\ 0 < \mu < 1$$

(两个集合 A 与 B 都是凸的), 就有

$$\mu z_1 + (1-\mu)z_2 = ((\mu x_1 + (1-\mu)x_2) - (\mu y_1 + (1-\mu)y_2)) \in C, \quad 对所有\ 0 < \mu < 1.$$

最后, 因为 $A \cap B = \varnothing$, 所以 $O \notin C$.

由 (ii), 存在非零的 $\ell \in X'$ 使得 $\ell(z) < \ell(0) = 0$ 对所有 $z \in C$ 成立, 或者等价地对任意的 $z = x - y$ 成立, 其中 $x \in A, y \in B$. 换言之,

$$\ell(x) < \ell(y), \quad 对所有\ x \in A\ 及所有\ y \in B.$$

因此存在 $\gamma \in \mathbb{R}$ 使得

$$\ell(x) \leqslant \gamma \leqslant \inf_{y \in B} \ell(y), \quad 对所有\ x \in A,$$

这就完成了证明. □

出现在上述证明中的次线性泛函 $p: X \to [0, \infty[$ 称为凸集 C 的 Minkowski *泛函*或*规范*或*支撑函数*.

对实赋范向量空间的两个子集 A 与 B, 如果存非零的连续线性泛函 $\ell \in X'$, 数 $\gamma \in \mathbb{R}$ 及 $\delta > 0$ 使得 (图 5.10-2)

$$\ell(x) \leqslant \gamma - \delta, \quad 对所有\ x \in A \quad 及 \quad \gamma + \delta \leqslant \ell(y), \quad 对所有\ y \in B,$$

则称其*严格地被超平面分离*.

下述定理给出实赋范向量空间的两个子集严格地被超平面分离的充分条件. 对于扩展到复空间的情况, 见习题 5.10-3(2).

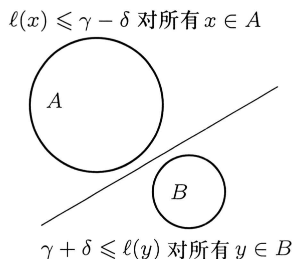


图 5.10-2　被超平面 (在此图中为一直线) 严格分离的 $\mathbb{R}^2$ 的两子集 A 与 B

定理 5.10-2 (Hahn-Banach 定理的第二几何形式: 凸集的严格分离) 设 A 与 K 为实赋范向量空间 X 的两个非空子集, 满足下述性质:

$$A \text{ 是凸且闭的;}\quad K \text{ 是凸且紧的;}\quad A \cap K = \varnothing.$$

那么一定存非零的 $\ell \in X', \gamma \in \mathbb{R}$ 及 $\delta > 0$ 使得

$$\ell(x) \leqslant \gamma - \delta < \gamma + \delta < \ell(y), \quad \text{对所有 } x \in A, y \in K.$$

证明　对任意 $r > 0$, 设

$$A(r) := \bigcup_{x \in A} B(x; r), \quad K(r) := \bigcup_{y \in K} B(y; r).$$

则集合 $A(r)$ 与 $K(r)$ 都是非空、凸而且开的. 此外, 存在 $r_0 > 0$ 使得 $A(r) \cap K(r) = \varnothing$ 对所有 $r \leqslant r_0$ 成立. 为说明这一结论, 假定其不成立, 则存在 $x_n \in A, v_n \in X, y_n \in K, w_n \in X, n \geqslant 1$, 使得

$$x_n + v_n = y_n + w_n, \quad \text{对所有 } n \geqslant 1,$$
$$v_n \to 0, w_n \to 0, \quad \text{当 } n \to \infty \text{ 时}.$$

集合 K 是紧的, 一定存在一个在 K 中收敛的子列 $(y_{\sigma(n)})_{n=1}^{\infty}$. 所以子列 $(x_{\sigma(n)})_{n=1}^{\infty}$ 也在 X 中收敛, 且 $\lim_{n\to\infty} x_\sigma(n) \in A$, 此因 A 是闭的. 因此 $\lim_{n\to\infty} x_{\sigma(n)} = \lim_{n\to\infty} y_{\sigma(n)} \in A \cap K$, 与假设矛盾.

所以, 根据定理 5.10-1, 两个集合 $A(r_0)$ 与 $B(r_0)$ 被一个超平面分离, 即存在非零的 $\ell \in X'$ 及 $\gamma \in \mathbb{R}$ 使得

$$\ell(x+v) \leqslant \gamma \leqslant \ell(y+w)$$

对所有 $x \in A, y \in K$ 及所有 $v, w \in X, \|w\| = r_0$ 成立. 因此

$$\begin{aligned}\ell(x) + r_0\|\ell\| &= \ell(x) + \sup_{\|v\|=r_0} \ell(v) \leqslant \gamma \\ &\leqslant \ell(y) + \inf_{\|w\|=r_0} \ell(w) = \ell(y) - r_0\|\ell\|, \quad \text{对所有 } x \in A, y \in B,\end{aligned}$$

证明完成 ($\delta := r_0\|\ell\| > 0$, 此因 $\ell \neq 0$).　□

习题

5.10-1 设 X 是维数 $\geqslant 2$ 的向量空间. X 中的一个超平面是 X 中形如 $\{x \in X; \ell(x) = 0\}$ 的任一子集, 其中 $\ell : X \to \mathbb{K}$ 是一非零线性泛函.

(1) 证明 X 的一个子空间 H 是一个超平面的充要条件是商空间 H/X 具有维数 1.

(2) 设 $\ell : X \to \mathbb{K}$ 及 $\widetilde{\ell} : X \to \mathbb{K}$ 是两个非零线性泛函. 证明 $\{x \in X; \ell(x) = 0\} = \{x \in X; \widetilde{\ell}(x) = 0\}$ 的充分必要条件是存在 $\alpha \neq 0$ 使得 $\ell = \alpha\widetilde{\ell}$.

5.10-2 设 X 是赋范向量空间. 证明: 一个线性泛函 $\ell : X \to \mathbb{K}$ 是连续的充分必要条件为超平面 $\{x \in X; \ell(x) = 0\}$ 是闭的.

5.10-3 (复向量空间中 Hahn-Banach 定理的几何形式)

(1) 在定理 5.10-1 的同样假设下, 只是现在 X 为复赋范向量空间. 证明, 存在非零的 $\ell \in X'$ 及 $\gamma \in \mathbb{R}$ 使得

$$\mathrm{Re}\,\ell(x) < \gamma \leqslant \mathrm{Re}\,\ell(y), \quad 对所有\ x \in A, y \in B.$$

(2) 在定理 5.10-2 的同样假设下, 只是现在 X 为复赋范向量空间. 证明, 存在非零的 $\ell \in X', \gamma \in \mathbb{R}$ 及 $\delta > 0$ 使得

$$\mathrm{Re}\,\ell(x) \leqslant \gamma - \delta < \gamma + \delta \leqslant \mathrm{Re}\,\ell(y), \quad 对所有\ x \in A, y \in K.$$

5.10-4 设 X 是一赋范向量空间, Y 是 X 的子空间. 证明 $\overline{Y} = X$ 的充分必要条件是, $\ell = 0$ 是满足下述性质的唯一的连续线性泛函 $\ell : \ell(y) = 0$ 对所有 $y \in Y$ 成立.

提示: 利用定理 5.10-2.

5.10-5 设 X 是赋范向量空间, 而 $f : X \to \mathbb{R}$ 是凸连续函数. 证明存在 $\ell \in X'$ 及 $c \in \mathbb{R}$ 使得 $f(x) > \ell(x) + c$ 对所有 $x \in X$ 成立.

5.11 对偶算子; Banach 闭值域定理

假设给定两个无穷维的赋范线性空间 X, Y 及一个映射 $A \in \mathcal{L}(X; Y)$. 我们希望确定, 给定任意向量 $y \in Y$, 是否存在 $x \in X$ 满足如下线性方程

$$Ax = y.$$

虽然这个方程解唯一性的问题, 即决定 $\mathrm{Ker}\,A = \{0\}$ 与否, 通常容易解决; 但存在性的问题, 即 $\mathrm{Im}\,A = Y$ 与否, 一般来说是相当困难的.

但是, 倒是可以给出使 $\mathrm{Im}\,A = Y$ 成立或刻画 $\mathrm{Im}\,A$ 的简单且非常有用的充分必要条件. 只是它们不是用算子 A 本身, 而是用 A 的对偶算子 A', 一个在下述定理 5.11-1 中定义的从 Y' 到 X' 的连续线性算子, 来刻画的.

这种算子是两个 Hilbert 空间之间线性连续算子的伴随算子到任意赋范向量空间上的自然推广: 设 X 及 Y 是两个 Hilbert 空间, 而 $\sigma : X' \to X$ 及 $\tau : Y' \to Y$ 是相应

的 F. Riesz 等距 (定理 4.6-1). 可以直接验证, 伴随算子 $A^* \in \mathcal{L}(Y;X)$ (如定理 4.7-2 中所定义) 与对偶算子 $A' \in \mathcal{L}(Y',X')$ (如下述定理 5.11-1 中所定义) 之间具有如下关系:

$$A' = \sigma^{-1}A^*\tau.$$

这些充分且必要的条件一起构成了奇妙的 Banach 闭值域定理, 该定理的命名源自以下事实, 一个向量空间在线性算子作用下的直接像也称为该线性算子的值域.

这个定理由两部分构成. 在第一部分中, 借助于 Y' 的子空间Ker A', 在 Im A 是闭的假设下, 给出 Y 的子空间Im A 特别有用的刻画, 见定理 5.11-5(a) 及 (c). 这个刻画在下面 Babuška-Brezzi 定理 (定理 6.12-1) 的证明中, 在 Stokes 方程组解的存在性 (定理 6.14-1) 以及 Donati 条件的充分性 (定理 6.19-4—定理 6.19-6) 的证明中起到关键作用.

在第二部分中, 特别给出了同样有用的使 Im $A = Y$ 的充分必要条件, 还是利用对偶算子 A', 即 A' 是闭映像 Im A' 的内射; 见定理 5.11-6.

这里顺便提一下, 这个定理的叙述很简单, 但其证明却不容易. 特别是, 这个证明本质上依赖 Banach 开映射定理, 赋范向量空间中的 Hahn-Banach 定理以及 Hahn-Banach 定理的几何表示.

在本节其余部分, X' 中的元素以典型的符号 x' (而不是 ℓ) 表示. 而为了简单起见, 在不至于引起混乱的情形下, 对诸如 $(A'y')(x)$ 等表达式将采用较简洁的符号 $A'y'(x)$ 等表示. 回忆一下, 赋范向量空间 X 的对偶空间 $X' := \mathcal{L}(X;\mathbb{K})$, 装备由 $\|x'\| = \sup_{x\neq 0}\dfrac{|x'(x)|}{\|x\|}$ 定义的范数, 是 Banach 空间.

定理 5.11-1 (对偶算子)　设 X 与 Y 是同一数域 $\mathbb{K}$ 上的两个赋范向量空间. 给定任意算子 $A \in \mathcal{L}(X;Y)$, 存在唯一的算子 $A' \in \mathcal{L}(Y';X')$, 称为 A 的对偶算子或简称 A 的对偶, 使得

$$A'y'(x) = y'(Ax), \quad \text{对所有 } x \in X \text{ 及所有 } y' \in Y'.$$

此外,

$$\|A'\|_{\mathcal{L}(Y';X')} = \|A\|_{\mathcal{L}(X;Y)}.$$

证明　(i) 给定 $y' \in Y' = \mathcal{L}(Y;\mathbb{K})$, 映射

$$A'y' : x \in X \to A'y'(x) := y'(Ax) \in \mathbb{K}$$

作为连续线性映射的复合, 是一个连续线性泛函.

(ii) 由这种方式定义的映射 $A' : Y' \to X'$ 是线性的, 此因对任意 $y', \widetilde{y}' \in Y'$,

$$\begin{aligned} A'(y' + \widetilde{y}')(x) &= (y' + \widetilde{y}')(Ax) = y'(Ax) + \widetilde{y}'(Ax) \\ &= A'y'(x) + A'\widetilde{y}'(x) \end{aligned}$$

对所有 $x \in X$ 成立; 而且对任意的 $\alpha \in \mathbb{K}$ 及 $y' \in Y'$, 有

$$\begin{aligned} A'(\alpha y')(x) &= (\alpha y')(Ax) = \alpha(y'(Ax)) \\ &= \alpha(A'y'(x)) = (\alpha A'y')(x) \end{aligned}$$

对所有 $x \in X$ 成立.

(iii) 线性算子 $A' : Y' \to X'$ 是连续的, 此因

$$\begin{aligned} \|A'y'\| &= \sup_{x \neq 0} \frac{|A'y'(x)|}{\|x\|} = \sup_{x \neq 0} \frac{|y'(Ax)|}{\|x\|} \\ &\leqslant \|A\|\|y'\|, \quad \text{对所有 } y' \in Y'. \end{aligned}$$

所以一方面有 $\|A'\| \leqslant \|A\|$. 另一方面由定理 5.9-5,

$$\begin{aligned} \|Ax\| &= \sup_{y' \neq 0} \frac{|y'(Ax)|}{\|y'\|} = \sup_{y' \neq 0} \frac{|A'y'(x)|}{\|y'\|} \\ &\leqslant \left(\sup_{y' \neq 0} \frac{\|A'y'\|}{\|y'\|} \right) \|x\| = \|A'\|\|x\|, \quad \text{对所有 } x \in X. \end{aligned}$$

故 $\|A\| \leqslant \|A'\|$. 所以 $\|A'\| = \|A\|$. □

注 在定理 6.14-1 中将给出一个无限维赋范向量中的有趣实例, 其中将证明算子

$$\begin{aligned} A : \mu \in L_0^2(\Omega) &:= \left\{ \mu \in L^2(\Omega); \int_\Omega \mu dx = 0 \right\} \\ &\to A\mu := \mathbf{grad}\, \mu \in \boldsymbol{H}^{-1}(\Omega) \end{aligned}$$

的对偶算子是

$$A' : \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \to A'\boldsymbol{v} := -\mathrm{div}\, \boldsymbol{v} \in L_0^2(\Omega).$$

虽然下面关于对偶算子紧性的充分条件并不是马上要用到的, 但为方便起见, 仍将其列于此 (其逆命题也成立; 见习题 5.11-1). 要注意的是, 或许想不到, Ascoli-Arzelà 定理在其证明中起着实质性的作用.

定理 5.11-2 设 X 与 Y 是两个实赋范向量空间, $A : X \to Y$ 是紧线性算子. 则对偶算子 $A' : Y' \to X'$ 也是紧的.

证明 设 $B := \overline{B_X(0;1)}$, 则 $K := \overline{A(B)}$ 是 Y 的紧子集 (因 A 是紧的). 特别地, 存在 M 使得 $K \subset \overline{B_Y(0;M)}$.

设 $(y'_n)_{n=1}^\infty$ 是 Y' 中任意的有界序列, 不失一般性可以假定其满足 $\|y'_n\|_{Y'} \leqslant 1, n \geqslant 1$. 则函数

$$f_n : y \in K \to f_n(y) := y'_n(y), \quad n \geqslant 1$$

形成空间 $\mathcal{C}(K)$ 中等度连续且有界族, 此因对每个 $n \geqslant 1$,

$$|f_n(y) - f_n(\widetilde{y})| = |y'_n(y - \widetilde{y})| \leqslant \|y - \widetilde{y}\|, \quad \text{对所有 } y, \widetilde{y} \in K,$$
$$\|f_n\|_{\mathcal{C}(K)} = \sup_{y \in K} |f_n(y)| \leqslant M.$$

所以, 由 Ascoli-Arzelà 定理 (定理 3.10-1), 存在一个在 $\mathcal{C}(K)$ 中收敛的子列 $(f_{\sigma(n)})_{n=1}^{\infty}$. 关系式

$$\begin{aligned}
\|A'y'_{\sigma(m)} - A'y'_{\sigma(n)}\|_{X'} &= \sup_{\|x\|_X \leqslant 1} |(A'(y'_{\sigma(m)} - y'_{\sigma(n)}))(x)| \\
&= \sup_{\|x\|_X \leqslant 1} |(y'_{\sigma(m)} - y'_{\sigma(n)})(Ax)| \\
&\leqslant \sup_{y \in K} |(y'_{\sigma(m)} - y'_{\sigma(n)})(y)| \\
&= \sup_{y \in K} |f_{\sigma(m)}(y) - f_{\sigma(n)}(y)| = \|f_{\sigma(m)} - f_{\sigma(n)}\|_{\mathcal{C}(K)}
\end{aligned}$$

说明序列 $(A'y'_{\sigma(n)})_{n=1}^{\infty}$ 是 X' 中的 Cauchy 序列, 由于 X' 是完备的 (是对偶空间), 故在 X' 中收敛. 所以 $A' : Y' \to X'$ 是紧的. □

作为寻求保证 $\text{Im } A = Y$ 的充分必要条件的第一步, 我们首先证明一个关于 A 的对偶 A' 的很简单的条件, 它等价于稍逊的要求 $\overline{\text{Im } A} = Y$. 注意, 如果 X 与 Y 都是 Hilbert 空间, 这个条件可立即由定理 4.7-2(b) 中确立的关系式 $Y = \text{Ker } A^* \oplus \overline{\text{Im } A}$ 得出.

定理 5.11-3 设 X 与 Y 是同一数域上的两个赋范向量空间, $A \in \mathcal{L}(X; Y)$. 则下述两个条件等价:

(a) 算子 A 有稠密的值域, 即 $\overline{\text{Im } A} = Y$.

(b) 对偶 $A' : Y' \to X'$ 是单射, 即 $\text{Ker } A' = \{0\}$.

证明 由定理 5.9-7, $\overline{\text{Im } A} = Y$ 的充分必要条件是

$$\{y' \in Y'; y'(z) = 0, \text{ 对所有 } z \in \text{Im } A\} = \{0\}.$$

所以定理叙述中的等价性可简单地由下述关系式得到:

$$\begin{aligned}
&\{y' \in Y'; y'(z) = 0, \text{ 对所有 } z \in \text{Im } A\} \\
&= \{y' \in Y'; y'(Ax) = 0, \text{ 对所有 } x \in X\} \\
&= \{y' \in Y'; A'y'(x) = 0, \text{ 对所有 } x \in X\} \\
&= \{y' \in Y'; A'y' = 0\} = \text{Ker } A'.
\end{aligned}$$

□

对于算子 $A \in \mathcal{L}(X; Y)$, 我们现在给出使其满足 $\text{Im } A = Y$ 的第一个基本的充分必要条件, 从而对本节开始提出的方程 $Ax = y$ 解存在性问题给出了第一个回答.

实际上, 这个结果可以归并入 Banach 闭值域定理的第二部分 (定理 5.11-6), 但在证明 Banach 闭值域定理的第一部分 (定理 5.11-5) 时, 要用到这一结果, 故需先行予以讨论.

定理 5.11-4 设 X 与 Y 是同一数域上的两个 Banach 空间, $A \in \mathcal{L}(X;Y)$. 则下面两个条件等价:

(a) 算子 $A: X \to Y$ 是满射, 即 $\operatorname{Im} A = Y$.

(b) 存在常数 $C > 0$ 使得对偶算子 $A': Y' \to X'$ 满足

$$\|y'\| \leqslant C\|A'y'\|, \quad \text{对所有 } y' \in Y'.$$

证明 对 $\mathbb{K} = \mathbb{R}$ 的情况给出证明, 当 $\mathbb{K} = \mathbb{C}$ 时的证明作为习题留给读者 (习题 5.11-3). 在下面的 (i) 中, 证明 (a) 可导出 (b). 在 (ii) 及 (iii) 部分, 证明 (b) 可导出 (a).

符号 $B_Z(z;r)$ 表示在空间 Z 中, 球心位于 z、半径为 r 的开球.

(i) 作为一个满射, 根据 Banach 开映射定理 (定理 5.6-1), A 是开的. 特别地可得, $B_X(0;1)$ 在 A 作用下的像是 Y 中包含点 $0 = A0$ 的开子集, 即存在 $\varepsilon > 0$ 使得

$$B_Y(0;s) \subset A(B_X(0;1)).$$

所以对任意 $y' \in Y'$, 有

$$\begin{aligned}\|A'y'\| &= \sup_{x\in B_X(0;1)} |A'y'(x)| = \sup_{x\in B_X(0;1)} |y'(AX)| \\ &= \sup_{y\in A(B_X(0;1))} |y'(y)| \geqslant \sup_{y\in B_Y(0;s)} |y'(y)| = s\|y'\|,\end{aligned}$$

因此 (b) 中的不等式对于 $C = s^{-1}$ 成立.

(ii) 假定 (b) 成立. 那么定有

$$B_Y(0;C^{-1}) \subset \overline{A(B_X(0;1))}.$$

选取任意点 $y_0 \in Y$, 使其不属于集合

$$Z := \overline{A(B_X(0;1))}.$$

因为 Z 是 Y 的闭凸子集, 而 $\{y_0\}$ 是 Y 中的紧凸子集, 它与 Z 的交是空集, Hahn-Banach 定理的第二几何形式 (定理 5.10-2) 说明, Z 与 $\{y_0\}$ 能被一个超平面严格分离. 这意味着存在非零的 $\widetilde{y}' \in Y'$ 及 $\delta > 0$ 使得

$$\widetilde{y}'(y) \leqslant \gamma - \delta < \gamma + \delta \leqslant \widetilde{y}'(y_0), \quad \text{对所有 } y \in Z,$$

而且 $\widetilde{y}'(y_0) > 0$, 此因 $0 \in Z, \widetilde{y}'(0) = 0$. 因此由上不等式得

$$\sup_{y\in Z} \widetilde{y}'(y) \leqslant \gamma - \delta < \widetilde{y}'(y_0).$$

注意到 $y \in A(B_X(0;1))$ 意味着 $-y \in A(B_X(0;1))$, 我们有

$$\sup_{y \in Z} |\widetilde{y}'(y)| = \sup_{y \in Z} \widetilde{y}'(y) \leqslant \gamma - \delta < \widetilde{y}'(y_0).$$

设

$$y_0' := (\sup_{y \in Z} |\widetilde{y}'(y)|)^{-1} \widetilde{y}', \quad 若 \ \sup_{y \in Z} |\widetilde{y}'(y)| > 0$$

或

$$y_0' := \frac{\delta}{\widetilde{y}'(y_0)} \widetilde{y}', \quad 对任意\ \delta > 1,\ 若\ \sup_{y \in Z} |\widetilde{y}'(y)| = 0,$$

这样, 我们找到一个非零的 $y_0' \in Y'$ 满足

$$|y_0'(y)| \leqslant 1, \quad 对所有\ y \in Z\ 且\ y_0'(y_0) > 1.$$

所以有

$$\|A'y_0'\| = \sup_{x \in B_X(0;1)} |A'y_0'(x)| = \sup_{x \in B_X(0;1)} |y_0'(Ax)| \leqslant \sup_{y \in Z} |y_0'(y)| \leqslant 1,$$

这就意味着

$$1 < y_0'(y_0) \leqslant \|y_0'\| \|y_0\| \leqslant C \|A'y_0'\| \|y_0\| \leqslant C \|y_0\|.$$

换言之, $y_0 \notin \overline{A(B_X(0;1))}$ 可推得 $\|y_0\| \geqslant C^{-1}$, 这就意味着

$$B_Y(0; C^{-1}) \subset \overline{A(B_X(0;1))}.$$

(iii) 假定 (b) 成立. 则存在 $s > 0$ 使得

$$B_Y(0; s) \subset A(B_X(0;1)).$$

所以算子 A 是满射.

细心的读者肯定会注意到, 上面 (ii) 部分中所证明的东西正是 Banach 开映射定理 (定理 5.6-1) 证明的 (ii) 部分里已经证明的. 只要逐字逐句地重述同一证明的部分 (iii)(幸好在那里满射的假设是不需要的), 即可推得存在 $s > 0$ 使得 $B_Y(0; s) \subset A(B_X(0;1))$.

由此显然可得, A 是满射. □

给定赋范向量空间 X 的子集 Z, 相应地有其对偶空间 X' 的子集 Z'. 由

$$Z^0 := \{x' \in X';\ x'(z) = 0,\ 对所有\ z \in Z\}$$

定义的 X' 的子空间及相应地, 由

$$^0(Z') := \{x \in X;\ z'(x) = 0,\ 对所有\ z' \in Z'\}$$

定义的 X 的子空间分别称为 Z 及 Z' 的极集; 有时也称为正交集 (反映它推广了内积空间中正交补的概念; 参阅 4.5 节) 或分别称其为 Z 及 Z' 的零化子 (这名称有点怪). 下面我们将会发现, 这些子空间将自然地出现在下个定理的刻画中 (也可参阅定理 5.11-3 或习题 5.11-2).

考虑到最终在定理 5.11-6(c) 中给出 A 是满射的第二个等价条件, 首先需要一个基础性的结果, 这就特别回答了本节开始提出的关于 $\text{Im } A$ 刻画的问题; 见下一个定理中的 (c).

要注意, 在 Hilbert 空间中当 $\text{Im } A$ 和 $\text{Im } A^*$ 是闭集时, 成立关系式 $\text{Im } A = (\text{Ker } A^*)^{\perp}$ 和 $\text{Im } A^* = (\text{Ker } A)^{\perp}$, 下一定理中的关系式 (c) 和 (d) 构成其在一般 Banach 空间的自然的推广.

定理 5.11-5 (Banach 闭值域定理[26]; 第一部分) 设 X 与 Y 是同一数域上的两个 Banach 空间, $A \in \mathcal{L}(X;Y)$. 则下述两个条件等价:

(a) 算子 $A : X \to Y$ 有闭值域, 即 $\text{Im } A$ 是 Y 中的闭集.

(b) 对偶算子 $A' : Y' \to X'$ 有闭值域, 即 $\text{Im } A'$ 是 X' 中的闭集.

(c) $\text{Im } A = {}^0(\text{Ker } A') = \{y \in Y; y'(y) = 0,$ 对所有 $y' \in \text{Ker } A'\}$.

(d) $\text{Im } A' = (\text{Ker } A)^0 = \{x' \in X'; x'(x) = 0,$ 对所有 $x \in \text{Ker } A\}$.

证明 (i) 关系 (a) 意味着关系 (b).

定义 Y 的子空间 $\widetilde{Y}$ 及映射 $\widetilde{A} \in \mathcal{L}(X;\widetilde{Y})$ 如下:

$$\widetilde{Y} := \text{Im } A \quad \text{和} \quad \widetilde{A} : x \in X \to \widetilde{A}x := Ax \in \widetilde{Y} \subset Y.$$

由假设知, $\text{Im } A = \text{Im } \widetilde{A} = \widetilde{Y}$ 是 Y 的闭子空间.

给定任意的 $y' \in Y'$, 注意到包含关系 $\widetilde{Y} \subset Y$, 对所有 $y \in Y$ 均成立的估计式 $|y'(y)| \leqslant \|y'\|\|y\|$ 说明, y' 由下式定义一个元素 $\widetilde{y}' \in Y' : \widetilde{y}'(\widetilde{y}) = y'(\widetilde{y})$, 对所有 $\widetilde{y} \in \widetilde{Y}$. 所以有

$$\begin{aligned} A'y'(x) &= y'(Ax) = \widetilde{y}'(Ax) = \widetilde{y}'(\widetilde{A}x) \\ &= \widetilde{A}'\widetilde{y}'(x), \quad \text{对所有 } x \in X, \end{aligned}$$

此因 $Ax \in \widetilde{Y}$ 对所有 $x \in X$ 成立. 这样, 我们就找到了 $y' \in Y'$ 使得 $A'y' = \widetilde{A}'\widetilde{y}'$.

给定任意的 $\widetilde{y}' \in \widetilde{Y}' = \mathcal{L}(\widetilde{Y};\mathbb{K})$, 根据赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1), 存在 $y' \in Y' = \mathcal{L}(Y;\mathbb{K})$ 使得

$$y'(\widetilde{y}) = \widetilde{y}'(\widetilde{y}), \quad \text{对所有 } \widetilde{y} \in \widetilde{Y}.$$

[26] S. Banach [1932]: Théorie des Opérateurs Linéaires. Monografje Matematyczne, Volume 1, Warsaw.

这就有

$$A'y'(x) = y'(Ax) = \widetilde{y}'(Ax) = \widetilde{y}'(\widetilde{A}x) = \widetilde{A}'\widetilde{y}'(x), \quad 对所有\ x \in X.$$

这样, 我们就找到了 $y' \in Y'$ 使得 $A'y' = \widetilde{A}'\widetilde{y}'$.

把上面得到的两种关系结合起来, 就得到, 在空间 X' 中成立

$$\mathrm{Im}\ A' = \mathrm{Im}\ \widetilde{A}'.$$

因为 $\widetilde{A} : X \to \widetilde{Y}$ 是满射, 而 $\widetilde{Y}$ (作为 Banach 空间的闭子集) 是完备的, 应用 Banach 开映射定理 (定理 5.6-1), 说明 A 将开集映射为开集. 特别地, 存在 $\delta > 0$ 使得 $B_{\widetilde{Y}}(0; 2\delta) \subset \widetilde{A}(B_X(0;1)) = A(B_X(0;1))$. 这个包含关系意味着, 给定任意的 $\widetilde{y} \in \widetilde{Y}, \|\widetilde{y}\| = \delta$, 一定存在 $x \in X$ 使得 $Ax = \widetilde{y}$ 且 $\|x\| < 1 = \dfrac{\|Ax\|}{\delta}$. 这样, 对任意的 $\widetilde{y}' \in \widetilde{Y}'$,

$$\begin{aligned}\|\widetilde{y}\| &= \frac{1}{\delta}\sup_{\|\widetilde{y}\|=\delta} |\widetilde{y}'(\widetilde{y})| \leqslant \frac{1}{\delta}\sup_{\|x\|<1} |\widetilde{y}'(\widetilde{A}x)| \\ &= \frac{1}{\delta}\sup_{\|x\|<1} |\widetilde{A}'\widetilde{y}'(x)| = \frac{1}{\delta}\|\widetilde{A'\widetilde{y}'}\|.\end{aligned}$$

这就证明了, 逆算子 $(\widetilde{A}')^{-1} : \mathrm{Im}\ \widetilde{A}' \subset X' \to \widetilde{Y}'$ 是有定义的, 从而 $\widetilde{A}' : \widetilde{Y}' \to \mathrm{Im}\ \widetilde{A}'$ 是双射并且是具有连续逆的连续线性算子.

因此 $\mathrm{Im}\ \widetilde{A}'$ 在 $\widetilde{X}'$ 中是闭的. 事实上, 设 $\widetilde{x}'_n \in \mathrm{Im}\ \widetilde{A}', n \geqslant 0$, 且 $\widetilde{x}'_n \to \widetilde{x}'$ 在 $\widetilde{X}'$ 中成立; 则 $(\widetilde{A}')^{-1}\widetilde{x}'_n$ 在 $\widetilde{Y}'$ 中收敛. 设 $\widetilde{y}' = \lim_{n\to\infty}(\widetilde{A}')^{-1}\widetilde{x}'_n$, 则 $\widetilde{x}'_n = \widetilde{A}'(\widetilde{A}')^{-1}\widetilde{x}'_n \to \widetilde{A}'\widetilde{y}'$, 由定义它属于 $\mathrm{Im}\ \widetilde{A}'$.

所以, 如所证, $\mathrm{Im}\ A' = \mathrm{Im}\ \widetilde{A}'$ 在 $\widetilde{X}'$ 中是闭的.

(ii) 关系 (b) 意味着关系 (a).

Y 的闭子空间 $\widehat{Y}$ 与算子 $\widehat{A} \in \mathcal{L}(X; \widehat{Y})$ 由下式定义:

$$\widehat{A} : x \in X \to \widetilde{A}x := Ax \in \widehat{Y} := \overline{\mathrm{Im}\ A} \subset Y.$$

因为根据 $\widehat{Y}$ 的定义, 空间 $\mathrm{Im}\ \widehat{A} = \mathrm{Im}\ A$ 在其中是稠密的, 故对偶算子 $\widehat{A}' : \widehat{Y}' \to X'$ 是单射 (定理 5.11-3). 将 (i) 中关于算子 $\widetilde{A}$ 的同样推导用于算子 $\widehat{A}$, 可以得到

$$\mathrm{Im}\ A' = \mathrm{Im}\ \widehat{A}', \quad 在空间\ X'\ 中.$$

由假设 $\mathrm{Im}\ A'$ 在 X' 中是闭的. 由于 X' 作为对偶空间是完备的, 所以 $\mathrm{Im}\ A'$ 是完备的, 因此 $\mathrm{Im}\ \widehat{A}'$ 也是完备的.

因为单射算子 $\widehat{A}' : \widehat{Y}' \to \mathrm{Im}\ \widehat{A}'$ (由其构造知) 是满射, 又可以利用 Banach 开映射定理 (定理 5.6-1) 证明, $(\widehat{A}')^{-1} : \mathrm{Im}\ \widehat{A}' \to \widehat{Y}'$ 是连续的. 这意味着存在一个常数 C 使得

$$\|\widehat{y}'\| \leqslant C\|\widehat{A}'\widehat{y}'\|, \quad 对所有\ \widehat{y}' \in \widehat{Y}'$$

(定理 2.9-4). 然后定理 5.11-4 说明, 算子 $\widehat{A}: X \to \widehat{Y}$ 是满射, 即

$$\operatorname{Im} \widehat{A} = \widehat{Y} = \overline{\operatorname{Im} A}.$$

但 $\operatorname{Im} \widehat{A} = \operatorname{Im} A$, 所以 $\operatorname{Im} A$ 是闭的. □

(iii) 关系 (a) 与关系 (c) 等价.

如果 $y \in \operatorname{Im} A$, 则对每个 $y' \in Y'$, $y'(y) = y'(Ax) = A'y'(x)$ 对某个 $x \in X$ 成立; 因此对所有的 $y' \in \operatorname{Ker} A'$ 有 $y'(y) = 0$. 这就证明了包含关系 $\operatorname{Im} A \subset {}^0(\operatorname{Ker} A')$ 总是成立的.

其次, 假设 $\operatorname{Im} A$ 是闭的, 但包含关系 ${}^0(\operatorname{Ker} A') \subset \operatorname{Im} A$ 不成立, 即存在 $y_0 \in {}^0(\operatorname{Ker} A')$ 使得 $y_0 \notin \operatorname{Im} A$. 那么由定理 5.9-6 (赋范向量空间中的 Hahn-Banach 定理的一个推论, 其在此适用是因为 $\operatorname{Im} A$ 是 Y 的闭子空间), 存在 $y_0' \in Y'$ 使得

$$y_0'(y_0) \neq 0 \quad \text{及} \quad y_0'(y) = 0, \quad \text{对所有 } y \in \operatorname{Im} A.$$

这就有 $A'y_0'(x) = y_0'(Ax) = 0$ 对所有 $x \in X$ 成立, 意味着 $y_0' \in \operatorname{Ker} A'$; 于是我们应该有 $y_0'(y_0) = 0$, 导致矛盾.

因此如果 $\operatorname{Im} A$ 是闭的, 我们就得到 $\operatorname{Im} A = {}^0(\operatorname{Ker} A')$, 即 (a) 意味着 (c). 而 (c) 意味着 (a) 是显然的, 这是因为极集总是闭的.

(iv) 关系 (b) 等价于关系 (d).

如果 $x' \in \operatorname{Im} A'$, 则对每个 $x \in X$, $x'(x) = A'y'(x) = y'(Ax)$ 对某个 $y' \in Y'$ 成立; 因此对所有 $x \in \operatorname{Ker} A$ 有 $x'(x) = 0$ 成立. 这就证明了包含关系 $\operatorname{Im} A' \subset (\operatorname{Ker} A)^0$ 总是成立的.

下面我们证明, 如果 $\operatorname{Im} A'$ 是闭的, 则有 $(\operatorname{Ker} A)^0 \subset \operatorname{Im} A'$. 为此, 定义商空间 $\dot{X} := X/\operatorname{Ker} A$, 进而定义双射连续线性算子 $\dot{A}: \dot{X} \to \operatorname{Im} A \subset Y$ 如下:

$$\dot{A}\dot{x} := A\widetilde{x}, \quad \text{对每个 } \dot{x} \in \dot{X},$$

其中 $\widetilde{x}$ 是 $\dot{x}$ 的任一元素. 装备以商范数, 商空间 $\dot{X}$ 就是一个 Banach 空间 (这是因为 $\operatorname{Ker} A$ 是闭子空间; 参阅定理 3.6-5), 而 $\operatorname{Im} A$ 也是一个 Banach 空间 (若 $\operatorname{Im} A'$ 是闭的, 由 (ii) 知 $\operatorname{Im} A$ 也是闭的). 根据 Banach 开映射定理 (定理 5.6-2) 的推论, $\dot{A}$ 的逆映射 $\dot{A}^{-1}: \operatorname{Im} A \to \dot{X}$ 也是连续线性算子.

给定任意元素 $x' \in (\operatorname{Ker} A)^0$, 设 $\dot{x}' \in \dot{X}'$ 由下式定义: $\dot{x}'(\dot{x}) := x'(\widetilde{x})$ 对每个 $\dot{x} \in \dot{X}$, 其中 $\widetilde{x}$ 是 $\dot{x}$ 的任一元素 (这个定义有意义, 这是因为如果 $x' \in (\operatorname{Ker} A)^0$ 则对所有 $x \in \operatorname{Ker} A$ 有 $x'(x) = 0$). 函数

$$y' := \dot{x}' \circ \dot{A}^{-1} : \operatorname{Im} A \to \mathbb{R}$$

因此是连续线性泛函, 即 $y' \in (\operatorname{Im} A)'$. 设 $\widetilde{y}' \in Y'$ 是 y' 的外延. 则对任意 $x \in X$, 有

$$\begin{aligned} A'\widetilde{y}'(x) &= \widetilde{y}'(Ax) = y'(Ax) = \dot{x}(\dot{A}^{-1}Ax) \\ &= \dot{x}'(\dot{A}^{-1}\dot{A}\dot{x}) = \dot{x}'(\dot{x}) = x'(x). \end{aligned}$$

所以

$$x' = A'\widetilde{y}' \in \mathrm{Im}\, A',$$

这就证明了 $(\mathrm{Ker}\, A)^0 \subset \mathrm{Im}\, A'$.

我们因此得到结论, 如果 $\mathrm{Im}\, A'$ 是闭的, 则有 $\mathrm{Im}\, A' = (\mathrm{Ker}\, A)^0$, 即 (b) 意味着 (d). 而 (d) 意味 (b) 则是显然的, 这是因为极集总是闭的. □

注 判定 $\mathrm{Im}\ A$ 在 Y 中是闭集的其他必要充分条件在习题 5.11-4 中给出.

借助于定理 5.11-5, 我们现在能够给出一个算子 A 满足 $\mathrm{Im}\ A = Y$ 的第二个基本充分必要条件 (见下定理 (c)), 从而对本节开始所提出的问题提供第二个回答. 对同一问题的第一个回答也被复述于下定理中 (见 (b)).

定理 5.11-6 (Banach 闭值域定理[27]; 第二部分) 设 X 与 Y 为在同一数域上的两个 Banach 空间, $A \in \mathcal{L}(X; Y)$. 则下述三个条件是等价的:

(a) 算子 $A : X \to Y$ 是满射, 即 $\mathrm{Im}\ A = Y$.

(b) 存在常数 C 使得对偶算子 $A' : Y' \to X'$ 满足

$$\|y'\| \leqslant C\|A'y'\|, \quad \text{对所有 } y' \in Y'.$$

(c) 对偶算子 A' 是单射且 $\mathrm{Im}\ A'$ 在 X' 中是闭的.

证明 (a) 与 (b) 之间的等价性已经在定理 5.11-4 中确立.

假定 (a) 成立. 则 $\mathrm{Im}\ A = \overline{\mathrm{Im}\ A} = Y$, 因此由定理 5.11-3 知 A' 是单射, 并且由定理 5.11-5, $\mathrm{Im}\ A'$ 是闭的. 所以 (a) 意味着 (c).

假定 (c) 成立, 则映射 (仍表示为) $A' : Y' \to \mathrm{Im}\ A'$ 是双射, 而且 $\mathrm{Im}\ A'$ 作为 X' 的闭子空间是完备的. 因此将 Banach 开映射定理的推论 (定理 5.6-2) 应用于 $A' \in \mathcal{L}(Y'; \mathrm{Im}\ A')$ 说明 $(A')^{-1} : \mathrm{Im}\ A' \to Y'$ 也是连续的. 这意味着存在常数 C 使得 $\|y'\| \leqslant C\|A'y'\|$ 对所有 $y' \in Y'$ 成立 (定理 2.9-4). 所以 (c) 意味着 (b). □

定理 5.11-6 中的条件 (b) 有时被用于线性边值问题的分析. 这是由于它断言, 为了确立一个表示为 $Ax = y$ 的偏微分方程解 x 的存在性 (在此, X 与 Y 是特定的函数空间, 其中 X 融合了某种边界条件; $A : X \to Y$ 是一个偏微分算子, 而 y 是方程的右端), 只需对对偶方程 $A'y' = x'$ 的任一给定的解 y' 得到一个形如 $\|y'\| \leqslant C\|x'\|$ 的先验界限, 其中 C 为与 x' 无关的常数. 特别值得注意的是, 并不需要验证对偶方程是否有解. 这种思路为确立此类问题解的存在性提供非常有力的技巧.

最后, 我们顺便提一下, Banach 闭值域定理对更一般的闭稠定线性算子也成立[28]; 然而, 这种推广在本书的其余部分中并不需要.

---

[27] S. BANACH [1932]: Théorie des Opérateurs Linéaires. Monografje Matematyczne, Volume 1, Warsaw.

[28] 参阅 YOSIDA [1965, 第 7 章第 5 节] 或 BREZIS [2001, 第 2.6 及 2.7 节].

习题

5.11-1 设 X 是实赋范向量空间, Y 是实 Banach 空间, 又设 $A \in \mathcal{L}(X;Y)$ 且其对偶算子 $A':Y' \to X'$ 是紧的. 证明 A 是紧的.

5.11-2 设 X 是赋范向量空间, 而 Z 是 X 的一个闭子集. 这样, 商空间 X/Z 也是一个赋范向量空间 (定理 2.2-3).

(1) 证明子空间 $Z^0 := \{x' \in X'; x'(z) = 0$, 对所有 $z \in Z\}$ 在 X' 中是闭的, 而且存在一个从 Y' 到 X'/Z^0 上的线性等距.

(2) 证明存在一个从 $(X/Z)'$ 到 Z^0 上的线性等距.

5.11-3 对 X 与 Y 是复 Banach 空间的情况, *逐字逐句地*证明定理 5.11-4 成立.

提示: 在定理 5.11-4 证明的 (ii), 利用关于复向量空间中凸集严格分离的 Hahn-Banach 定理的几何形式 (习题 5.10-3(2)), 证明存在 $y_0' \in \widetilde{Y}'$ 使得 $|y_0'(y)| \leqslant 1$ 对所有 $y \in C$ 成立, 并且 $|y'(y_0)| > 1$.

5.11-4 设 X 与 Y 是 Banach 空间, 而 $A \in \mathcal{L}(X;Y)$.

(1) 假定 A 是单射. 证明 $\mathrm{Im}\,A$ 在 Y 中是闭集的必要充分条件是, 存在常数 C 使得 $\|x\| \leqslant C\|Ax\|$ 对所有 $x \in X$ 成立.

(2) 假定 A 不是单射. 证明 $\mathrm{Im}\,A$ 在 Y 中是闭集的必要充分条件是, 存在常数 C 使得 $\|[x]\| \leqslant C\|Ax\|$ 对所有 $x \in X$ 成立, 其中 $\|[x]\|$ 表示在商空间 $X/\mathrm{Ker}\,A$ (它也是 Banach 空间, 见定理 3.6-5) 中 $[x]$ 的范数.

5.11-5 设 $(X, (\cdot,\cdot))$ 是 Hilbert 空间, 而 $A \in \mathcal{L}(X)$ 是对称算子. 证明, 如果存在一个常数 $\alpha > 0$ 使得 $(Ax, x) \geqslant \alpha\|x\|^2$ 对所有 $x \in X$ 成立, 则 $A: X \to X$ 是单射且 $\mathrm{Im}\,A = X$ (因此, 对每一个 $y \in X$, 方程 $Ax = y$ 有且只有一个解 $x \in X$).

5.12 弱收敛和弱 * 收敛

在*有限维*赋范向量空间中, 每一个*有界序列一定包含一个收敛的子序列* (有限维空间中有界子集的闭包是紧的; 见定理 2.7-1(c)). 但这一结论在*无限维*空间中不一定成立. 例如, 在内积空间中 (4.8 节) 一个可列无限的标准正交族 (e_i) 有界, 对所有的 i, $\|e_i\| = 1$, 但不包含任何收敛子列, 此因如果 $i \neq j$, 则 $\|e_i - e_j\| = \sqrt{2}$.

这样, 一个自然的问题就产生了: 是否有*另一种收敛*概念, 它在*无限维*赋范向量空间中具有类似的性质? 结果发现, 合适的概念是*弱收敛*[29], 其定义将在下面给出. 由于下面将证明 (定理 5.14-4), *在自反 Banach 空间中的任一有界序列一定包含一个弱收敛的子列* (所谓自反 Banach 空间系指在特定的线性等距意义下, 能与其对偶的对偶等同的空间; 参见 5.14 节), 因此这一概念给出上面所提问题一个正面的回答. 特别要注意, 这一性质将被用于强制泛函极小化序列的处理, 所以它在变分学中 (第 9 章) 起着关键的作用.

[29] 这个概念是 David Hilbert 在 1906 年前后引入的.

设 X 是赋范向量空间, X' 表示其对偶. 一个元素为 $x_n \in X$ 的序列 $(x_n)_{n=1}^{\infty}$ 称为在 X 中是弱收敛的, 如果存在 $x \in X$ 使得

$$\text{对每个 } x' \in X', \quad x'(x_n) \to x'(x), \quad \text{当 } n \to \infty \text{ 时},$$

而这个 x 则称为序列 $(x_n)_{n=1}^{\infty}$ 的弱极限. 弱收敛用"半箭头" $\rightharpoonup$, 即由

$$x_n \rightharpoonup x, \quad \text{当 } n \to \infty \text{ 时}$$

表示, 即可与强收敛加以区分, 后者以"全箭头" $\to$, 即由下式表示:

$$x_n \to x, \quad \text{当 } x \to \infty \text{ 时}.$$

事实上"强收敛"只是按范数拓扑收敛的别称: 它意味着当 $n \to \infty$ 时, $\|x_n - x\| \to 0$.

我们注意到, 在 Hilbert 空间 $(X, (\cdot,\cdot))$ 中, 序列 $(x_n)_{n=1}^{\infty}$ 弱收敛于 x 当且仅当对每个 $y \in X$, 当 $n \to \infty$ 时 $(x_n, y) \to (x, y)$. 此因根据 F. Riesz 等距, X 的对偶 X' 等同于 X.

例如, 由下式定义的序列 $(f_n)_{n=1}^{\infty}$

$$f_n(\theta) := \sin n\theta, \quad 0 \leqslant \theta \leqslant 2\pi$$

在 Hilbert 空间 $L^2(0, 2\pi)$ 中弱收敛于 0. 为了说明这一点, 我们知道, 给定任意函数 $g \in L^2(0, 2\pi)$, 数

$$a_n := \frac{1}{\pi} \int_0^{2\pi} g(\varphi) \cos n\varphi d\varphi, \quad n \geqslant 0 \quad \text{及}$$
$$b_n := \frac{1}{\pi} \int_0^{2\pi} g(\varphi) \sin n\varphi d\varphi, \quad n \geqslant 1$$

是 g 的"经典" Fourier 级数的系数. 这样, 由 Parseval 公式 (定理 4.9-2), 有

$$\frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} |a_n|^2 + \sum_{n=1}^{\infty} |b_n|^2 = \frac{1}{\pi} \|g\|_{L^2(0,2\pi)}^2 < \infty,$$

所以 $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = 0$. 这也说明由 $g_n(\theta) = \cos n\theta, 0 \leqslant \theta \leqslant 2\pi$, 定义的序列 $(g_n)_{n=1}^{\infty}$ 当 $n \to \infty$ 时在 $L^2(0, 2\pi)$ 中也弱收敛于 0.

但是, 序列 $(f_n)_{n=1}^{\infty}$ (就此而言, 序列 $(g_n)_{n=1}^{\infty}$ 也如此) 在 $L^2(0, 2\pi)$ 中不强收敛. 事实上, 它甚至不包含任何强收敛的子列, 这是因为 $\left(\frac{1}{\sqrt{\pi}} f_n\right)_{n=1}^{\infty}$ 是一标准正交族.

下定理给出弱收敛与强收敛之间两个简单的关系. 更深入详尽的关系将在稍后给出 (定理 5.12-3 与 5.13-1).

定理 5.12-1 (a) 在任一赋范向量空间里, 若当 $n \to \infty$ 时, $x_n \to x$, 则当 $n \to \infty$ 时, $x_n \rightharpoonup x$.

(b) 在有限维赋范向量空间中, 任何弱收敛序列也强收敛. 故这两个收敛概念在有限维空间中是一致的.

证明 设 $(X, \|\cdot\|)$ 是赋范向量空间, 且当 $n \to \infty$ 时, $x_n \to x$ 在 X 中成立. 则

$$\text{对每个 } x' \in X', \quad |x'(x_n) - x'(x)| \leqslant \|x'\| \|x_n - x\|,$$

这就证明了 (a).

设 X 是有限维的, $(e_i)_{i=1}^k$ 是 X 的基, 又设 X 装备了范数 $\|\cdot\|_1$ (使证明更简单). 每一线性泛函

$$x_j' : x = \sum_{i=1}^{k} x_i e_i \in X \to x_j \in \mathbb{K}, \quad 1 \leqslant j \leqslant k$$

显然都是连续的, 此因 $|x_j| \leqslant \|x\|$ 对所有 $x \in X$ 成立. 弱收敛 $x_n = \sum_{i=1}^k x_i^n e_i \rightharpoonup x = \sum_{i=1}^k x_i e_i$ 意味着对每个 $1 \leqslant j \leqslant k, x_j^n \to x_j$. 因此当 $n \to \infty$ 时 $\|x_n - x\|_1 \to 0$, 这就证明了 (b). □

对于弱收敛, 立即会想到两个自然的问题: 弱收敛的极限是唯一的吗? 弱收敛序列是有界的吗? 令人感到意外的是, 要 (正面地) 回答这两个看似平常的问题绝不是平凡的事. 事实上, 要回答第一个问题 (对强收敛序列答案是显然的) 至少需要赋范向量空间中的 Hahn-Banach 定理 (因此也就需要选择公理); 而对第二个问题的回答 (对强收敛序列也是显然的) 至少需要赋范向量空间中的 Hahn-Banach 定理以及 Banach-Steinhaus 定理 (因此既需要选择公理也需要 Baire 定理). 这些性质将在下一定理中确立 (见 (a) 与 (b)), 该定理也对弱极限的范数给出了一个上界 (见 (c)).

定理 5.12-2 在任一赋范向量空间中, 下述性质均成立:

(a) 弱收敛序列的极限是唯一的.

(b) 弱收敛序列是有界的.

(c) 设当 $n \to \infty$ 时, $x_n \rightharpoonup x$, 则

$$\|x\| \leqslant \liminf_{n \to \infty} \|x_n\|.$$

证明 设 $(x_n)_{n=1}^{\infty}$ 是赋范向量空间 $(X, \|\cdot\|)$ 中的弱收敛序列, 而 $x, \widetilde{x} \in X$ 使得

$$\text{对每个 } x' \in X', \quad x'(x) = x'(\widetilde{x}) = \lim_{n \to \infty} x'(x_n).$$

这就有 $x'(x - \widetilde{x}) = 0$ 对所有 $x' \in X'$ 成立, 所以由定理 5.9-5 (Hahn-Banach 定理的一个推论), $x = \widetilde{x}$. (a) 得证.

对于每个 $n \geqslant 1$, 由下式定义映射 $J_n : X' \to \mathbb{K}$:

$$J_n : x' \in X' \to J_n(x') := x'(x_n),$$

它显然是线性的. 而且 $J_n \in (X')' := \mathcal{L}(X'; \mathbb{K})$, 此因由定理 5.9-5 有

$$\|J_n\|_{(X')'} = \sup_{\left\{\substack{x' \in X' \\ x' \neq 0}\right.} \frac{|J_n(x')|}{\|x'\|} = \sup_{\left\{\substack{x' \in X' \\ x' \neq 0}\right.} \frac{|x'(x_n)|}{\|x'\|} = \|x_n\|.$$

此外, 由于对每个 $x' \in X'$,

$$\lim_{n\to\infty} J_n(x') = x'(x),$$

其中 x 表示序列 $(x_n)_{n=1}^{\infty}$ 的弱极限, 故序列 $(J_n(x'))_{n=1}^{\infty}$ 在 $\mathbb{K}$ 中收敛. 设线性映射 $J_x : X' \to \mathbb{K}$ 由下式定义:

$$J_x : x' \in X' \to J_x(x') := \lim_{n\to\infty} J_n(x') = x'(x).$$

仍由定理 5.9-5 有

$$\|J_x\|_{(X')'} = \sup_{x'\neq 0} \frac{|x'(x)|}{\|x'\|} = \|x\|,$$

所以它是连续的.

因为空间 X' (作为一个对偶空间, 见定理 3.2-3) 是完备的, 应用 Banach-Steinhaus 定理的推论 (定理 5.3-2), 可证得

$$\sup_{n\geqslant 1} \|J_n\|_{(X')'} < \infty \quad 及 \quad \|J_x\|_{(X')'} \leqslant \liminf_{n\to\infty} \|J_n\|_{(X')'},$$

它们正是

$$\sup_{n\geqslant 1} \|x_n\| < \infty \quad 及 \quad \|x\| \leqslant \liminf_{n\to\infty} \|x_n\|,$$

所以 (b) 与 (c) 得证. □

注　在 Hilbert 空间 $(X, (\cdot,\cdot))$ 中, 弱极限的唯一性是很容易证明的. 这是由于如果 $x_n \underset{n\to\infty}{\rightharpoonup} x$ 及 $x_n \underset{n\to\infty}{\rightharpoonup} \widetilde{x}$, 这意味着 $(x - \widetilde{x}, y) = 0$ 对所有 $y \in X$ 成立. 这立即导出 $x = \widetilde{x}$.

顺便提一下, 为什么由 $f_n(\theta) = \sin n\theta, 0 \leqslant \theta \leqslant 2\pi$, 定义的序列 $(f_n)_{n=1}^{\infty}$ 在 $L^2(0, 2\pi)$ 中不强收敛? 弱极限的唯一性对此给出了另一个理由. 这是因为, 如果存在 $f \in L^2(0, 2\pi)$ 使得 $f_n \underset{n\to\infty}{\longrightarrow} f$, 这意味着 $f_n \underset{n\to\infty}{\rightharpoonup} f$. 根据弱收敛序列极限的唯一性 (定理 5.12-2(a)), 就有 $f = 0$. 但这是不可能的, 因为简单的计算就说明 $\|f_n\|_{L^2(0,2\pi)} = \sqrt{\pi} \neq 0$ 对所有 $n \geqslant 1$ 成立.

我们现在给出一个使弱收敛序列强收敛的非常有用的充分条件. 回忆一下 (2.17 节), 一个赋范向量空间 $(X, \|\cdot\|)$ 称为一致凸的, 如果给定任意的 $\varepsilon > 0$, 存在 $\delta(\varepsilon) > 0$ 使得

$$\|x\| = \|y\| = 1 \text{ 且 } \|x - y\| \geqslant \varepsilon \text{ 意味着 } \left\|\frac{x+y}{2}\right\| \leqslant 1 - \delta(\varepsilon).$$

而且 $\ell^p, L^p(\Omega), 1 < p < \infty$, 以及任一内积空间都是一致凸的 (习题 2.17-8, 2.17-9 及定理 4.1-2).

定理 5.12-3　*设 X 是一致凸的赋范向量空间, 而元素为 $x_n \in X$ 的序列 $(x_n)_{n=1}^{\infty}$ 以及 $x \in X$ 满足*

$$x_n \rightharpoonup x \text{ 且 } \|x_n\| \to \|x\|, \quad 当\ n \to \infty\ 时.$$

则当 $n \to \infty$ 时, $x_n \to x$.

证明 如果 $x = 0$, 结果显然成立. 如果 $x \neq 0$, 则存在 $n_0 \geqslant 1$ 使得 $x_n \neq 0$ 对所有 $n \geqslant n_0$ 成立, 此因当 $n \to \infty$ 时 $\|x_n\| \to \|x\| > 0$. 令

$$y := \frac{x}{\|x\|}, \quad y_n = \frac{x_n}{\|x_n\|}, \quad \text{对所有 } n \geqslant n_0.$$

这样, 对所有 $x' \in X'$ 及 $n \geqslant n_0$, 有

$$x'(y_n) = \frac{1}{\|x_n\|}x'(x_n) \to \frac{1}{\|x\|}x'(x) = x'(y), \quad \text{当 } n \to \infty \text{ 时}.$$

故 $y_n \rightharpoonup y$, 因此当 $n \to \infty$ 时 $(y_n + y) \rightharpoonup 2y$.

而定理 5.12-2(c) 说明

$$\begin{aligned} 2 = \|2y\| &\leqslant \liminf_{n\to\infty} \|y_n + y\| \\ &\leqslant \limsup_{n\to\infty} \|y_n + y\| \leqslant 2. \end{aligned}$$

又因 $\|y_n + y\| \leqslant \|y_n\| + \|y\| = 2$, 上式给出

$$\left\|\frac{y_n + y}{2}\right\| \to 1, \quad \text{当 } n \to \infty \text{ 时}.$$

而这个关系又意味着

$$\|y_n - y\| \to 0, \quad \text{当 } n \to \infty \text{ 时}.$$

如若不然, 则存在 $\varepsilon > 0$ 及一个子序列 $(y_{\sigma(n)})_{n=n_0}^{\infty}$ 使得 $\|y_{\sigma(n)} - y\| \geqslant \varepsilon$ 对所有 $n \geqslant n_0$ 成立. 但一致凸的假设却意味着 $\left\|\dfrac{y_{\sigma(n)} + y}{2}\right\| \leqslant 1 - \delta(\varepsilon)$ 对某个 $\delta(\varepsilon) > 0$ 成立, 与前面结果矛盾.

下述关系

$$x_n - x = \|x_n\|y_n - \|x\|y = \|x_n\|(y_n - y) + (\|x_n\| - \|x\|)y,$$

结合序列 $(x_n)_{n=1}^{\infty}$ 的有界性 (定理 5.12-2(b)), 就证明了当 $n \to \infty$ 时 $\|x_n - x\| \to 0$. □

注 当一致凸空间 X 是 Hilbert 空间 $(X, (\cdot,\cdot))$ 时, 有一个非常简单的证明, 只要在下式中取当 $n \to \infty$ 时的极限即可:

$$\|x_n - x\|^2 = \|x_n\|^2 - 2(x, x_n) + \|x\|^2.$$

根据 F. Riesz 表示定理, 在此, 当 $n \to \infty$ 时弱收敛 $x_n \rightharpoonup x$ 意味着当 $n \to \infty$ 时 $(x, x_n) \to (x, x) = \|x\|^2$.

下面我们讨论将线性或双线性算子作用于弱收敛序列的效果 (双线性算子及相应的空间如 $\mathcal{L}_2(X \times Y; \mathbb{K})$ 在 2.11 节中定义).

定理 5.12-4 设 X 及 Y 为同一数域 $\mathbb{K}$ 上的赋范向量空间.

(a) 设 $A \in \mathcal{L}(X; Y)$. 则

$$x_n \underset{n\to\infty}{\rightharpoonup} x \text{ 在 } X \text{ 中意味着 } Ax_n \underset{n\to\infty}{\rightharpoonup} Ax \text{ 在 } Y \text{ 中}.$$

(b) 设 $A \in \mathcal{L}(X; Y)$ 是紧的. 则

$$x_n \underset{n\to\infty}{\rightharpoonup} x \text{ 在 } X \text{ 中意味着 } Ax_n \underset{n\to\infty}{\rightarrow} Ax \text{ 在 } Y \text{ 中}.$$

(c) 设 $B \in \mathcal{L}_2(X \times Y; \mathbb{K})$. 则

$$x_n \underset{n\to\infty}{\rightharpoonup} x \text{ 在 } X \text{ 中而 } y_n \underset{n\to\infty}{\rightarrow} y \text{ 在 } Y \text{ 中意味着 } B(x_n, y_n) \rightarrow B(x, y) \text{ 在 } \mathbb{K} \text{ 中}.$$

证明 (i) 设 $A' \in \mathcal{L}(Y'; X')$ 表示 $A \in \mathcal{L}(X; Y)$ 的对偶 (定理 5.11-1). 则 $A'y' \in X'$ 对所有 $y' \in Y'$ 成立, 因此由弱收敛的定义,

$$\begin{aligned} y'(Ax_n) = A'y'(x_n) &\underset{n\to\infty}{\rightarrow} A'y'(x) \\ &= y'(Ax), \quad \text{对所有 } y' \in Y'. \end{aligned}$$

这就证明了 (a).

(ii) 设 $A \in \mathcal{L}(X; Y)$ 是紧算子, 而 $x_n \underset{n\to\infty}{\rightharpoonup} x$ 在 X 中成立. 因为弱收敛序列是有界的 (定理 5.12-2(b)), 序列 $(Ax_n)_{n=1}^{\infty}$ 包含一个在 Y 中强收敛的子列 $(Ax_{\sigma(n)})_{n=1}^{\infty}$ (根据紧算子的定义, 见 2.10 节). 进而根据 (a), 其极限为 Ax (强收敛序列也弱收敛, 见定理 5.12-1(a)). 由于这个极限是唯一的, 故整个序列 $(Ax_n)_{n=1}^{\infty}$ 强收敛于 Ax. 这证明了 (b).

(iii) (c) 的证明可由下面关系式

$$B(x_n, y_n) - B(x, y) = B(x_n, y_n - y) + B(x_n - x, y),$$

并结合以下结果得到: 弱收敛序列 $(x_n)_{n=1}^{\infty}$ 的有界性, B 的连续性 (这意味着, 对每一个 $y \in Y$, 映射 $x \in X \rightarrow B(x, y) \in \mathbb{K}$ 是 X 上的连续线性泛函) 以及弱和强收敛的定义. □

应当指出的是, 如果两个序列仅仅是弱收敛, 性质 (c) 并不一定成立. 作为例子, 考察连续双线性形式

$$B : (f, g) \in L^2(0, 2\pi) \times L^2(0, 2\pi) \rightarrow B(f, g) := \int_0^{2\pi} f(\theta)g(\theta)d\theta,$$

以及由 $f_k(\theta) = \sin k\theta, 0 \leqslant \theta \leqslant 2\pi$, 定义的序列 $(f_k)_{k=1}^{\infty}$, 我们在前面已知它在 $L^2(0, 2\pi)$ 中弱收敛于零. 但因为 $B(f_k, f_k) = \pi$ 对所有 $k \geqslant 1$ 成立, 故序列 $(B(f_k, f_k))_{k=1}^{\infty}$ 不收敛于 $B(0, 0) = 0$.

我们以不加证明地叙述关于弱收敛的若干重要补充结论来结束这一节.

设 $(X, \|\cdot\|)$ 是赋范向量空间. 其对偶空间由所有的线性形式 $x' : X \to \mathbb{R}$ 组成, 当 X 装备以由其范数 $\|\cdot\|$ 诱导的拓扑, 也称为 X 上的强拓扑时, 这些线性形式是连续的.

然而, 同一空间 X 也可以装备其弱拓扑, 它被定义为 X 上使对偶空间 X' 的所有元素 x', 作为从装备了这种拓扑的 X 到 $\mathbb{K}$ 的函数, 保持连续的 "最弱" 拓扑 (定理 1.7-8 保证了这种拓扑的存在). 在此 "最弱" 一词的含义是, X 上任何具有同样性质的其他拓扑都包含更多的开集. X 的一子集对弱拓扑是开的, 对强拓扑也是开的, 但反之不然.

可以证明下述关于弱拓扑的一些基本性质[30).

定理 5.12-5 设 X 是赋范向量空间.

(a) 如果空间 X 是有限维的, 则强弱拓扑等同. 所以, 在此种情况下, 弱拓扑是可以赋范的.

(b) 如果空间 X 是无限维的, 则存在关于强拓扑的开集, 它们关于弱拓扑不是开的. 进而, 在此种情况下, 弱拓扑是不能尺度化的.

(c) X 上的弱拓扑是 Hausdorff 拓扑 (1.6 节).

(d) 元素为 $x_n \in X$ 的序列 $(x_n)_{n=1}^{\infty}$ 关于 X 的弱拓扑收敛于 $x \in X$ 的充分必要条件是 $x'(x_n) \to x'(x)$ 对所有 $x' \in X'$ 成立, 即其充分必要条件是序列 $(x_n)_{n=1}^{\infty}$ 弱收敛于 x.

上述定理中的 (b) 与 (d) 两部分说明为何无限维空间 "通常" 包含弱收敛但不强收敛的序列. 然而, 存在 "病态的" 空间, 如空间 ℓ^1, 其每一个弱收敛序列也都强收敛[31)!

性质 (d) 也说明, 弱收敛 (按照本节开始给出的定义) 正是相应于弱拓扑的收敛性.

顺便提一下, 判定一个拓扑能否根据收敛序列来定义是一个微妙的问题[32). 例如, 在空间 ℓ^1 中 (如前面指出的) 强收敛和弱收敛的序列是等同的; 但它们分别相应于强

[30) 其证明以及更多的性质, 可参阅 BREZIS [2011] 中 3.2 和 3.3 节的阐述.

[31) 这个结果构成 Schur 引理, 其证明可在 KESAVAN [2009, 5.1 节] 中找到.

[32) 关于这方面, 可参阅例如:

J. KISYNSKI [1959]: Convergence du type L. Colloquium Mathematicum **7**, 205–211.

S. P. FRANKLIN [1965]: Spaces in which sequences suffice. Fundamenta Mathematicae **57**, 107–115.

S. P. FRANKLIN [1967]: Spaces in which sequences suffice. Fundamenta Mathematicae **61**, 51–56.

R. M. DUDLEY [1964]: On sequential convergence. Transactions of the American Mathematical Society **112**, 483–507.

B. KRIPKE [1967]: One more reason why sequences are not enough. American Mathematical Monthly **74**, 563–565.

拓扑和弱拓扑, 因为 ℓ^1 是无限维空间, 两种拓扑肯定不同 (定理 5.12-5(b)).

另一个 "弱" 收敛的概念, 可由交换 X 和 X' 在弱收敛定义中的作用给出: 设 X 是赋范向量空间, X' 是其对偶. 一个由元素 $x'_n \in X'$ 组成的序列 $(x'_n)_{n=1}^{\infty}$ 称为在 X' 中是弱 * 收敛的, 如果存在 $x' \in X'$ 使得

$$\text{对每个 } x \in X, \quad x'_n(x) \to x'(x), \quad \text{当 } n \to \infty \text{ 时}.$$

这样的 x' 称为序列 $(x'_n)_{x=1}^{\infty}$ 的弱 * 极限, 它显然是唯一的. 弱 * 收敛以 "上面带星的半箭头" 表示, 即表示为

$$x'_n \overset{*}{\rightharpoonup} x', \quad \text{当 } n \to \infty \text{ 时}.$$

可以想见, 弱 * 收敛和弱收敛具有类似的性质; 见习题 5.12-4 和 5.14-6.

这样, 在对偶空间 X' 中可以定义三种不同类型的收敛: 强收敛:

$$x'_n \to x', \quad \text{当 } n \to \infty \text{ 时},$$

它意味着当 $n \to \infty$ 时 $\|x'_n - x'\|_{X'} \to 0$; 弱收敛:

$$x'_n \rightharpoonup x', \quad \text{当 } n \to \infty \text{ 时}.$$

它意味着对每个 $x'' \in (X')'$, 当 $n \to \infty$ 时 $x''(x'_n) \to x''(x')$; 以及上面所定义的弱 * 收敛.

以后 (5.14 节) 我们将会看到, 任一赋范向量空间 X 都可凭借特定的线性等距 $J : X \to J(X) \subset (X')'$, 等同于 $(X')'$ 的一个子空间. 换言之, 弱 * 收敛可视为弱收敛 (如本节开始所定义的) 的局限形式, 它只涉及 X' 的对偶空间 $(X')'$ 中那些属于映像 $J(X) \subset (X')'$ 的元素. 所以, 如果空间 X 使得等距 $J : X \to (X')'$ 是满射, 这两个概念就完全一致了. 这种空间称为自反的, 将在 5.14 节中予以讨论.

正如赋范向量空间 X 中的弱收敛那样, 弱 * 收敛相应于 X' 上的一个拓扑. 更明确地说, X' 上的弱 * 拓扑被定义为 X' 上使得所有映射 $\varphi_x : x' \in X' \to \varphi_x(x') := x'(x) \in \mathbb{K}, x \in X$ 连续的最弱拓扑 (定理 1.7-8)[33].

这样, 我们就可以确立下述结果 (与定理 5.12-5 相比较).

定理 5.12-6 设 X 是赋范向量空间.

(a) X' 的一个子集若对于 X' 的弱 * 拓扑是开的, 对 X' 的弱拓扑也是开的, 但反之不一定成立.

(b) X' 上的弱 * 拓扑是 Hausdorff 拓扑.

(c) X' 中的闭单位球对于 X' 的弱 * 拓扑是紧的.

(d) 元素 $x'_n \in X'$ 组成的序列 $(x'_n)_{n=0}^{\infty}$ 关于 X' 的弱 * 拓扑收敛于 $x' \in X'$ 的充分必要条件是它弱 * 收敛于 x'.

[33] 关于弱 * 拓扑极具可读性的讨论, 可参阅 BREZIS [2001, 3.4 节].

或许引入弱 * 拓扑最重要的理由是上述的性质 (c): 如果 X 是无限维的, 则 X' 的闭单位球关于 X' 的强拓扑就不可能是紧的 (由 F. Riesz 定理; 见定理 2.7-3), 但这同一个闭单位球关于 X' 的弱 * 拓扑却总是紧的 (即使 X 是无限维的).

习题

5.12-1 这个习题给出了一个关于弱收敛的充分条件. 设 X 是赋范向量空间, Y' 是其对偶 X' 的一个稠密子集. 又设元素 $x_n \in X$ 组成的序列 $(x_n)_{n=1}^{\infty}$ 及 $x \in X$ 满足

$$\sup_{n \geqslant 1} \|x_n\| < \infty \text{ 且 } y'(x_n) \underset{n\to\infty}{\to} y'(x), \quad \text{对每个 } y' \in Y'.$$

证明当 $n \to \infty$ 时 $x_n \rightharpoonup x$ (注意, 由定理 5.12-2(b), 条件 $\sup_{n\geqslant 1} \|x_n\| < \infty$ 也是必要的).

5.12-2 设 Ω 是 $\mathbb{R}^n$ 的开子集, $1 < p < \infty$, 又设函数 $f_k \in L^p(\Omega)$ 组成的序列 $(f_k)_{k=1}^{\infty}$ 是有界的且点态几乎处处收敛于 $f \in L^p(\Omega)$. 证明在 $L^p(\Omega)$ 中当 $k \to \infty$ 时 $f_k \rightharpoonup f$.

5.12-3 对每个整数 $n \geqslant 1$, 设函数 $f_n \in L^2(0,1)$ 定义如下:

$$f_n(x) := 0, \text{ 若 } \frac{j}{n} \leqslant x < \frac{j}{n} + \frac{1}{2n}, \quad 0 \leqslant j \leqslant n-1,$$
$$f_n(x) := 1, \text{ 若 } \frac{j}{n} + \frac{1}{2n} \leqslant x < \frac{j+1}{n}, \quad 0 \leqslant j \leqslant n-1.$$

(1) 证明序列 $(f_n)_{n=1}^{\infty}$ 在空间 $L^2(0,1)$ 中弱收敛.

提示: 利用习题 5.12-1.

(2) 序列 $(f_n)_{n=1}^{\infty}$ 在 $L^2(0,1)$ 中是否强收敛?

5.12-4 设 X 是 Banach 空间, $(x'_n)_{n=0}^{\infty}$ 是元素 x'_n 组成的在 X' 中弱 * 收敛的序列.

(1) 证明序列 $(x'_n)_{n=0}^{\infty}$ 在 X' 中是有界的.

(2) 证明 $(x'_n)_{n=0}^{\infty}$ 的弱 * 极限 x' 满足

$$\|x'\|_{X'} \leqslant \liminf_{n\to\infty} \|x'_n\|_{X'}.$$

注 这个结果构成定理 5.12-2 的 "弱 * 类似".

5.13 Banach-Saks-Mazur 定理

我们在 5.12 节中已看到, 由 $f_k(\theta) = \sin k\theta, 0 \leqslant \theta \leqslant 2\pi$, 定义的序列 $(f_k)_{k=1}^{\infty}$ 在空间 $L^2(0,2\pi)$ 中弱收敛于 0. 但在此空间中不强收敛. 然而, 存在函数 f_k 的凸组合 (2.16 节) 序列, 它在 $L^2(0,2\pi)$ 中强收敛于同一极限 0. 例如, 由以下诸式定义的序列 $(f_n)_{n=1}^{\infty}$ 及 $(h_n)_{n=1}^{\infty}$ 就如此 (习题 5.13-1):

$$f_n := \frac{1}{n}\sum_{k=1}^{n} f_k \quad \text{或} \quad h_n := \frac{1}{pn+1}\sum_{k=n}^{n+pn} f_k, \quad \text{对任一固定整数 } p \geqslant 1.$$

实际上, 这样的一个结论在任何赋范向量空间中都成立, 这可由下面近乎完美又非常有用的定理得知. 这一结果, 特别地是在变分学 (第 9 章) 中起着关键的作用.

其证明的 (i) 表明, 这一定理的证明主要依赖凸集被超平面分离的问题, 而这正是 Hahn-Banach 定理的第一几何形式所给出的 (定理 5.10-1).

要注意, 在下面定理的 (a) 和 (c) 两部分中, 标为 $(y_n)_{n=1}^{\infty}$ 和 $(z_n)_{n=1}^{\infty}$ 的序列强收敛于同一弱极限 x, 它们并不是全不能确定的凸组合序列. 而是如上面的例子那样, 第 n 个凸组合 y_n 正是给定弱收敛序列 $(x_n)_{n=1}^{\infty}$ 的前 n 项的一个组合, 而第 n 个凸组合 z_n 则是以 $(x_n)_{n=1}^{\infty}$ 的第 n 项作为开始的.

也还要注意, (c) 的证明基于另一个弱收敛和集合 C 凸性之间的关系, 即性质 (b), 当然其本身也是很重要的. 相比之下, (c) 的结论对任何强收敛序列, 不管闭集合 C 凸与否, 都成立.

定理 5.13-1 (Banach-Saks-Mazur 定理[34]) 设 X 是赋范向量空间.

(a) 设 $(x_k)_{k=1}^{\infty}$ 是 X 中的序列满足

$$x_k \rightharpoonup x, \quad \text{当 } k \to \infty \text{ 时}.$$

则对每个 $n \geqslant 1$, 存在 $\lambda_k^n \geqslant 0, 1 \leqslant k \leqslant n$, 满足 $\sum_{k=1}^{n} \lambda_k^n = 1$, 使得

$$y_n := \sum_{k=1}^{n} \lambda_k^n x_k \to x, \quad \text{当 } n \to \infty \text{ 时}.$$

(b) 设 C 是 X 的一个非空、凸闭子集, 而 $(x_k)_{k=1}^{\infty}$ 是点 $x_k \in C$ 的序列, 当 $k \to \infty$ 时该序列弱收敛于 $x \in X$. 则弱极限 x 属于 C.

(c) 设 $(x_k)_{k=1}^{\infty}$ 是 X 中的序列, 满足

$$x_k \rightharpoonup x, \quad \text{当 } k \to \infty \text{ 时}.$$

则对每个 $n \geqslant 1$, 存在一个整数 $m(n) \geqslant 0$ 及 $\mu_k^n \geqslant 0, n \leqslant k \leqslant n + m(n)$, 满足 $\sum_{k=n}^{n+m(n)} \mu_k^n = 1$, 使得

$$z_n := \sum_{k=n}^{n+m(n)} \mu_k^n x_k \to x, \quad \text{当 } n \to \infty \text{ 时}.$$

证明　回忆一下, co A 表示向量空间中一子集 A 的凸包 (2.16 节).

---

[34] S. Banach; S. Saks [1930]: Sur la convergence forte dans le champ L^p. Studia Mathematica **2**, 51–57.

S. Mazur [1933]: Über konvexe Mengen in linearen normierten Räumen. Studia Mathematica **5**, 70–84.

(i) (a) 的证明. 定义凸集

$$A_n := \mathrm{co}\left(\bigcup_{k=1}^{n}\{x_k\}\right), \quad \text{对每个整数 } n \geqslant 1 \text{ 及}$$
$$A := \bigcup_{n=1}^{\infty} A_n$$

(由其构造知, 每个集合 A_n 都是凸的; 又因 $A_n \subset A_{n+1}$ 对所有 $n \geqslant 1$ 成立, 故 A 也是凸的). 我们断言

$$\rho_n := \inf_{w \in A_n} \|x - w\| \to 0, \quad \text{当 } 0 \to \infty \text{ 时}.$$

如果断言不成立, 则存在 $\rho > 0$ 使得 $A \cap B(x;\rho) = \varnothing$ (此因 $\rho_n \geqslant \rho_{n+1}$ 对所有 $n \geqslant 1$ 成立). 因此, 根据 Hahn-Banach 定理的第一几何形式 (定理 5.10-1) 存在一个超平面分离 A 与 $B(x;\rho)$ (球 $B(x;\rho)$ 是开而且凸的): 这意味着存在非零的 $\ell \in X'$ 及 $\gamma \in \mathbb{R}$ 使得

$$\ell(x + \rho v) = \ell(x) + \rho\ell(v) \leqslant \gamma \leqslant \ell(w),$$
$$\text{对所有 } \|v\| = 1 \text{ 及所有 } w \in A.$$

这样就有

$$\ell(x) + \rho\|\ell\| = \ell(x) + \rho \sup_{\|v\|<1} \ell(v) \leqslant \ell(w), \quad \text{对所有 } w \in A.$$

在这个不等式中令 $w = x_n$, 使

$$\ell(x) + \rho\|\ell\| \leqslant \ell(x_n), \quad \text{对所有 } n \geqslant 1.$$

上式不可能成立, 因为根据弱收敛的定义, 当 $n \to \infty$ 时 $\ell(x_n) \to \ell(x)$.

所以当 $n \to \infty$ 时 $\inf_{w \in A_n} \|x - w\| \to 0$. 这意味着, 对每个 $n \geqslant 1$, 存在 $y_n \in A_n$ 使得当 $n \to \infty$ 时 $\inf_{w \in A_n} \|x - w\| = \|x - y_n\| \to 0$ (每个集 A_n 都是紧的). 等价地, 对每个 $n \geqslant 1$, 存在 $\lambda_k^n \geqslant 0, 1 \leqslant k \leqslant n$, 满足 $\sum_{k=1}^{n} \lambda_k^n = 1$ 使得

$$\left\| x - \sum_{k=1}^{n} \lambda_k^n x_k \right\| \to 0, \quad \text{当 } n \to \infty \text{ 时}$$

(显然不失一般性, 这里假定了所有 $x_1, x_2, \cdots, x_n$ 都进入每个凸组合 y_n).

(ii) (b) 的证明. 因为 C 是凸的, (a) 中给出的凸组合 $y_n = \sum_{k=1}^{n} \lambda_k^n x_k$ 属于 C. 此外, 还是由 (a), 当 $n \to \infty$ 时 $y_n \to x$. 由于 C 关于强拓扑是闭的, 所以 $x \in C$.

(iii) (c) 的证明. 定义凸集

$$C_n := \mathrm{co}\left(\bigcup_{k=n}^{\infty}\{x_k\}\right), \quad \text{对每个整数 } n \geqslant 1.$$

对每个 $n \geqslant 1, x_m \in C_n \subset \overline{C}_n$ 对所有 $m \geqslant n$ 成立, 而且 $(x_m)_{m=n}^{\infty}$ 弱收敛于 x; 所以由 (ii) 知 $x \in \overline{C}_n$ (每一个闭包 $\overline{C}_n$ 也都是凸的). 所以, 对每个 $n \geqslant 1$, 存在 $z_n \in C_n$ 使得 $\|x - z_n\| \leqslant \dfrac{1}{n}$; 等价地, 对每个 $n \geqslant 1$, 存在一个整数 $m(n) \geqslant 0$ 和 $\mu_k^n \geqslant 0, n \leqslant k \leqslant n + m(n)$, 满足 $\sum_{k=n}^{n+m(n)} \mu_k^n = 1$, 使得

$$\left\| x - \sum_{k=n}^{n+m(n)} \mu_k^n x_k \right\| \to 0, \quad \text{当 } n \to \infty \text{ 时}. \qquad \square$$

注 如果 X 是 Hilbert 空间 (在这种情况, X 可等同于其对偶空间), (b) 的证明变得非常简单, 无须依赖 Banach-Saks-Mazur 定理: 设 $(\cdot,\cdot)$ 表示 X 中的内积, 而 $P: X \to C$ 表示 X 到 C 上的投影算子, 因此满足 $(Px - x, y - Px) \geqslant 0$ 对所有 $y \in C$ 成立 (定理 4.3-1). 特别地, $(Px - x, x_k - Px) \geqslant 0$ 对所有 $k \geqslant 1$ 成立. 所以

$$-\|x - Px\|^2 = (Px - x, x - Px) = \lim_{k \to \infty} (Px - x, x_k - Px) \geqslant 0.$$

这就得 $x = Px \in C$.

Banach-Saks-Mazur 定理说明, *如果赋范向量空间 X 的凸子集关于 X 的强拓扑是闭的, 那么一定是序列弱闭的*, 即它包含其所有弱收敛序列的弱极限. 实际上, 可进而证明[35], 这种子集其实是"*弱闭的*", 即它在*关于 X 的弱拓扑意义下是闭的* (5.12 节).

习题

5.13-1 设 $f_k(\theta) := \sin k\theta, 0 \leqslant \theta \leqslant 2\pi, k \geqslant 1$. 证明序列 $\left(\dfrac{1}{n}\sum_{k=1}^{n} f_k\right)_{n=1}^{\infty}$ 和 $\left(\dfrac{1}{pn+1}\sum_{k=n}^{n+pn} f_k\right)$, 其中 $p \geqslant 1$ 为任意固定的整数, 在 $L^2(0, 2\pi)$ 中强收敛于 0.

5.14 自反空间; Banach-Eberlein-Šmulian 定理

设 X 是赋范向量空间. 则

$$X'' := (X')'$$

表示 X 的*双对偶空间*, 或简称 X 的*双对偶*, 即 X 的对偶空间的对偶空间. 作为一个对偶空间, 空间 X'' 当然是一个 Banach 空间, 其任一元素 $x'' \in X''$ 的范数由下式给出:

$$\|x''\|_{X''} = \sup_{x' \neq 0} \frac{|x''(x')|}{\|x'\|_{X'}}.$$

35) 例如, 可参阅 Brezis [1983, 定理 3.7].

我们将看到, 本节的基本结果 (Banach-Eberlein-Šmulian 定理; 见定理 5.14-4) 断言, 在 Banach 空间 X 中的任一有界序列均有弱收敛子列的充分必要条件是, X 能通过某一线性等距与其双对偶空间 X'' 等同. 为了正确地定义这一概念, 我们首先展示任一赋范向量空间如何通过自然的方式等同于其双对偶空间的一个子空间.

定理 5.14-1 设 X 为赋范向量空间. 对于每个 $x \in X$ 由下式:

$$Jx(x') := x'(x), \quad 对所有\ x' \in X'$$

定义的映射

$$J : x \in X \to Jx \in X''$$

是线性等距, 称为从 X 到 X'' 内的标准等距.

证明 给定任意元素 $x \in X$, 泛函

$$Jx : x' \in X' \to Jx(x') := x'(x) \in \mathbb{K}$$

是线性且连续的: 首先, 对所有 $\alpha, \beta \in \mathbb{K}$ 及所有 $x', y' \in X'$, 由于 X' 也是向量空间, 故

$$\begin{aligned} Jx(\alpha x' + \beta y') &= (\alpha x' + \beta y')(x) \\ &= \alpha x'(x) + \beta y'(x) = \alpha Jx(x') + \beta Jx(y'); \end{aligned}$$

其次 $Jx \in X''$, 此因

$$|Jx(x')| = |x'(x)| \leqslant \|x'\| \|x\|, \quad 对所有\ x' \in X'.$$

以这种方式定义的映射 $J : X \to X''$ 是一个等距, 这是因为由定理 5.9-5 (赋范向量空间中的 Hahn-Banach 定理的推论), 有

$$\begin{aligned} \|Jx\|_{X''} &= \sup_{x' \neq 0} \frac{|Jx(x')|}{\|x'\|} = \sup_{x' \neq 0} \frac{|x'(x)|}{\|x'\|} \\ &= \|x\|, \quad 对每个\ x \in X. \end{aligned}$$

□

注 出现在定理 5.12-2 证明中的映射 $J_n, n \geqslant 1$, 及 J_x 正是上面定理证明中的线性泛函 $Jx \in X''$ 的特殊情况.

一个赋范向量空间 X 要成为自反的, 只要在定理 5.14-1 中定义的标准等距 $J : X \to X''$ 是满射即可, 这样 X 就与其双对偶空间 X'' 等同. 换言之, 如果给定任意的 $x'' \in X''$, 存在 (唯一的) $x \in X$ 使得

$$x''(x') = x'(x), \quad 对所有\ x' \in X',$$

则 X 是自反的.

对于这个定义, 重要的是 X 与 X'' 的等同是用标准等距来实现的. 若不然, 就存在不是自反的 Banach 空间, 但能够通过一个线性等距与其双对偶等同[36].

注意, 作为一个对偶空间, 自反空间必定是完备的.

下面两个定理提供了自反 Banach 空间的实例. 更多的例子可见习题 5.14-1 及 5.14-2. 在习题 5.14-5 中给出一个反例, 装备以 sup 范数的空间 $\mathcal{C}[0,1]$.

定理 5.14-2 下述 Banach 空间是自反的:

(a) 有限维赋范向量空间;

(b) Hilbert 空间;

(c) 自反 Banach 空间的任一闭子空间;

(d) 自反 Banach 空间的对偶空间;

(e) 空间 $\ell^p, 1 < p < \infty$, 及 Lebesgue 空间 $L^p(\Omega), 1 < p < \infty$, 其中 Ω 是 $\mathbb{R}^n$ 的任一开子集.

证明 (i) (a) 的证明. 设 X 是有限维赋范向量空间. 给定 X 中的一组基 $(e_i)_{i=1}^n$, 关系式 $e_j'(e_i) = \delta_{ij}, 1 \leqslant i,j \leqslant n$, 定义 X' 中的一组基 $(e_j')_{j=1}^n$. 由此同样可利用关系式 $e_k''(e_j') = \delta_{jk}, 1 \leqslant j,k \leqslant n$, 定义 X'' 中的一组基 $(e_k'')_{k=1}^n$. 这样, 就可以直接验证对这种情况, 定理 5.14-1 中的标准等距 $J : X \to X''$ 由下式给出:

$$J\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i e_i'', \quad \text{对所有 } x = \sum_{i=1}^n x_i e_i \in X.$$

因此 J 是满射.

(ii) (b) 的证明. 给定 Hilbert 空间 $(X, (\cdot,\cdot)_X)$, 设 $\sigma : X' \to X$ 表示相应的 F. Riesz 等距. 则空间 X' 装备以由下式定义的内积 $(\cdot,\cdot)_{X'}$:

$$(x', y')_{X'} := \overline{(\sigma x', \sigma y')_X}, \quad \text{对每组 } x', y' \in X',$$

也是 Hilbert 空间 (定理 4.6-1). 设 $\sigma' : X'' \to X'$ 是相应的 F. Riesz 等距. 则可直接验证, 在这种情况标准等距 $J : X \to X''$ 由下式给出:

$$J = (\sigma \circ \sigma')^{-1}$$

(若 $\mathbb{K} = \mathbb{R}$, 是两个线性等距的复合; 若 $\mathbb{K} = \mathbb{C}$, 是两个半线性等距的复合, 都是一个线性等距).

(iii) (c) 的证明. 设 Y 是自反 Banach 空间的闭子空间. 我们要证明, 给定任意的 $y'' \in Y''$, 存在 $y \in Y$ 使得

$$y''(y') = y'(y), \quad \text{对所有 } y' \in Y'.$$

[36] R.C. James [1951]: A non-reflexive Banach space isometric with its second conjugate space. Proceedings of the National Academy of Sciences, USA **37**, 174–177.

设 $y'' \in Y''$ 是给定的. 则线性泛函

$$x'' : x' \in X' \to x''(x') := y''(x'|_Y) \in \mathbb{K}$$

是连续的, 其中 $x'|_Y$ 表示 x' 在 Y 上的限制, 此因

$$|y''(x'|_Y)| \leqslant \|y''\|\|x'|_Y\| \leqslant \|y''\|\|x'\|, \quad 对所有\ x' \in X'.$$

所以 $x'' \in X''$, 而且因为由假定 X 是自反的, 故存在 $y \in X$ 使得

$$x'(y) = x''(x') = y''(x'|_Y), \quad 对所有\ x' \in X'.$$

特别地可得, $x'(y) = 0$ 对所有那些其限制 $x'|_Y$ 为零的 $x' \in X'$ 成立. 由假设, Y 是闭的, 故 $y \in Y$ (如果 $y \notin Y$, 根据定理 5.9-6, 一定存在 $x' \in X'$ 使得 $x'|_Y = 0$ 但 $x'(y) \neq 0$).

给定任意的 $x' \in Y' = \mathcal{L}(Y; \mathbb{K})$, 设 $x' \in X' = \mathcal{L}(X; \mathbb{K})$ 是 y' 的任一延拓 (根据赋范向量空间中的 Hahn-Banach 定理, 这样延拓是存在的; 见定理 5.9-1). 则有

$$y''(y') = y''(x'|_Y) = x'(y) = y'(y)$$

(此因 $y' = x'|_Y$ 及 $y \in Y$), 这就是所要证明的.

(iv) (d) 的证明. 给定一个自反的 Banach 空间, 设 $J : X \to X''$ 表示 X 到其双对偶空间 X'' 上的标准等距, 而 $J' : X' \to (X')''$ 表示 X' 到其双对偶空间 $(X')''$ 上的标准等距.

给定任意 $x''' \in (X')''$, 定义映射

$$x' : x \in X \to x'(x) := x'''(Jx) \in \mathbb{K}.$$

要注意, 这个定义是有意义的, 这是因为 $Jx \in X''$ 而且

$$(X')'' = \mathcal{L}(\mathcal{L}(\mathcal{L}(X; \mathbb{K}); \mathbb{K}); \mathbb{K}) = (X'')'.$$

则定有 $x' \in X'$, 此因

$$|x'(x)| \leqslant \|x'''\|_{(X')''}\|Jx\|_{X''} = \|x'''\|_{(X')'''}\|x\|_X, \quad 对所有\ x \in X.$$

此外, x' 和 J 的定义以及 $J : X \to X''$ 的双射性质给出

$$x'''(x'') = x'''(Jx) = x'(x) = Jx(x') = x''(x'), \quad 对所有\ x'' = Jx \in X'',$$

由 J' 的定义, 这说明

$$x''' = J'x'.$$

因此 J' 是满射, 这就是所要证明的.

(v) (e) 的证明. 空间 $\ell^p, 1 < p < \infty$, 的自反性可从关于其对偶空间的刻画 (定理 3.5-1) 得出. 同样, 空间 $L^p(\Omega), 1 < p < \infty$, 可从 $L^p(\Omega)$ 中的 F. Riesz 表示定理 (定理 3.5-3) 得到. 当然, 对于 $p = 2$ 情况的自反性, 也可从 (b) 导出. □

确立空间 ℓ^p 和 $L^p(\Omega), 1 < p < \infty$, 自反性的另一个不同的途径是, 先证明它们是一致凸的 (习题 2.17-8 及 2.17-9), 然后利用下述关于自反性的重要充分条件. 鉴于这个结论的重要性, 将其列于下.

定理 5.14-3 (Milman-Pettis 定理[37]) 一致凸的 Banach 空间一定是自反的.

注 当然存在自反的 Banach 空间, 它们不是严格凸的, 更别说一致凸了. 例如, 空间 $(\mathbb{R}^n, \|\cdot\|_1)$ 和 $(\mathbb{R}^n, \|\cdot\|_\infty)$.

我们用线性泛函分析中最基本的定理之一来结束本章. 尤其要注意的是, (a) 的结果在确立强制且序列弱下半连续的泛函极小元素存在性 (定理 9.3-1) 中起着基本作用. (b) 则为证明一个空间是自反的, 提供了一个有效的途径; 例如, 可用来证明 Sobolev 空间 $W^{m,p}(\Omega), 1 < p < \infty$, 是自反的 (习题 6.11-2).

定理 5.14-4 (Banach-Eberlein-Šmulian 定理[38])

(a) 在自反 Banach 空间中, 任何有界序列一定包含一个弱收敛的子列.

(b) 反之, 如果一个 Banach 空间中每一个有界序列均包含一个弱收敛子列, 该空间一定是自反的[39].

证明 我们在空间 X 是可分的[40] 附加假设下 (下面所碰到的大部分函数空间都满足这一假设) 证明 (a).

(i) X 是自反的这一假设意味存在一个从 X 到 X'' 上的线性等距 (即标准等距); 所以与 X 一样, X'' 也是可分的. 又因为, 根据定义 $X'' = (X')'$, 故由定理 5.9-8 知, 空间 X' 也是可分的. 那么设 $x'_k \in X', k \geqslant 1$, 使得

$$X' = \overline{\bigcup_{k=1}^{\infty} \{x'_k\}}.$$

[37] 这一结果由以下独立给出:

D. P. Milman [1938]: On some criteria for the regularity of spaces of type (B). Doklady Akademii Nauk SSSR **20**, 243–246 (俄文).

B. J. Pettis [1939]: A proof that every uniformly convex space is reflexive. Duke Mathematical Journal **5**, 249–253.

关于这个结果的证明, 也可参阅 Yosida [1966, 第 5 章, 第 2 节], Diestel [1975] 或 Brezis [2011, 定理 3.31].

[38] (a) (在可分性的补充假设下) 的证明属于 Banach [1932]. (b) 归于:

V. L. Šmulian [1940]: Über lineare topologische Räume. Mathematiceskii Sbornik. N. S. **49**, 425–448.

W. F. Eberlein [1947]: Weak compactness in Banach spaces I. Proceedings of the National Academy of Sciences, USA **33**, 51–53.

[39] (b) 的证明可参阅 Yosida [1966, 第 5 章附录] 或 Dunford & Schwartz [1958, 第 5 章, 第 6 节].

[40] 不可分情形的证明参阅 Yosida [1966, 第 5 章附录] 或

R. Whitley [1967]: An elementary proof of the Eberlein-Šmulian theorem. Mathematische Annalen **172**, 116–118.

(ii) 设 $(x_n)_{n=1}^{\infty}$ 是元素 $x_n \in X$ 的有界序列. 故对每个 $x' \in X'$,

$$|x'(x_n)| \leqslant M\|x'\|, \text{ 对所有 } n \geqslant 1, \quad \text{其中 } M := \sup\|x_n\| < \infty.$$

序列 $(x'(x_n))_{n=1}^{\infty}$ 在 $\mathbb{K}$ 中是有界的, 包含一个收敛子列.

特别地, 序列 $(x_1'(x_n))_{n=1}^{\infty}$ 包含一个收敛子列 $(x_1'(x_{\sigma_1(n)}))_{n=1}^{\infty}$; 序列 $(x_2'(x_{\sigma_1(n)}))_{n=1}^{\infty}$ 同样在 $\mathbb{K}$ 中有界, 同样包含一个收敛子列 $(x_2'(x_{\sigma_2(n)}))_{n=1}^{\infty}$; 等等. 考察"对角线"序列

$$(x_{\sigma(n)})_{n=1}^{\infty}, \quad \text{其中 } \sigma(n) := \sigma_n(n), \quad n \geqslant 1.$$

由构造知, $(x_{\sigma(n)})_{n=1}^{\infty}$ 是序列 $(x_n)_{n=1}^{\infty}$ 的一个子列, 所以对每个整数 $k \geqslant 1$, 当 $n \to \infty$ 时序列 $(x_k'(x_{\sigma(n)}))_{n=1}^{\infty}$ 在 $\mathbb{K}$ 中收敛.

(iii) 我们下面证明, 实际上对每个 $x' \in X'$, 当 $n \to \infty$ 时序列 $(x'(x_{\sigma(n)}))_{n=1}^{\infty}$ 都在 $\mathbb{K}$ 中收敛.

设 $x' \in X'$ 及 $\varepsilon > 0$ 是给定的. 由 (i) 知, 存在整数 $k = k(x', \varepsilon) \geqslant 1$ 使得 $\|x' - x_k'\| \leqslant \dfrac{\varepsilon}{4M}$. 这样, 对任意整数 $m, n \geqslant 1$, 有

$$\begin{aligned} &|x'(x_{\sigma(m)}) - x'(x_{\sigma(n)})| \\ &\leqslant |x_k'(x_{\sigma(m)}) - x_k'(x_{\sigma(n)})| + |(x' - x_k')(x_{\sigma(m)} - x_{\sigma(n)})| \\ &\leqslant |x_k'(x_{\sigma(m)}) - x_k'(x_{\sigma(n)})| + \frac{\varepsilon}{2}, \end{aligned}$$

此因 $\|x_{\sigma(m)} - x_{\sigma(n)}\| \leqslant 2M$. 但作为收敛序列, $(x_k'(x_{\sigma(n)}))_{n=1}^{\infty}$ 是 Cauchy 序列, 故当 m 及 n 充分大时, 可使 $|x_k'(x_{\sigma(m)}) - x_k'(x_{\sigma(n)})|$ 充分小. 所以存在整数 $n_0 = n_0(k) = n_0(x', \varepsilon) \geqslant 1$ 使得

$$|x'(x_{\sigma(m)}) - x'(x_{\sigma(n)})| \leqslant \varepsilon, \quad \text{对所有 } m, n \geqslant n_0,$$

这说明 $(x'(x_{\sigma(n)}))_{n=1}^{\infty}$ 也是 Cauchy 序列. 所以 $(x'(x_{\sigma(n)}))_{n=1}^{\infty}$ 在 $\mathbb{K}$ 中收敛.

(iv) 设 $J : X \to X''$ 表示定理 5.14-1 中给出的线性等距, 因 X 被假定是自反的, 故它是满射. 对每个 $n \geqslant 1$, 连续线性泛函 $Jx_{\sigma(n)} \in X'' = \mathcal{L}(X'; \mathcal{K})$ 应由下式定义:

$$Jx_{\sigma(n)}(x') = x'(x_{\sigma(n)}), \quad \text{对所有 } x' \in X'.$$

由 (iii), 它具有下述性质: 对每个 $x' \in X'$, 当 $n \to \infty$ 时

$$\lim_{n\to\infty} Jx_{\sigma(n)}(x') \text{ 在 } \mathbb{K} \text{ 中存在.}$$

由于空间 X' 是完备的, 利用 Banach-Steinhaus 定理的推论 (定理 5.3-2) 知, 存在 $x'' \in X'' = \mathcal{L}(X'; \mathbb{K})$, 使得当 $n \to \infty$ 时

$$Jx_{\sigma(n)}(x') \to x''(x'), \quad \text{对每个 } x' \in X'.$$

但这又可表示为, 当 $n \to \infty$ 时

$$x'(x_{\sigma(n)}) \to x'(x), \quad \text{对每个 } x' \in X', \text{其中 } x := J^{-1}x''.$$

所以当 $n \to \infty$ 时, 子序列 $(x_{\sigma(n)})_{n=1}^{\infty}$ 弱收敛于 x. □

习题

5.14-1 证明, 如果 Banach 空间 X 的对偶是自反的, 则 X 本身也是自反的.

注 这个结果与定理 5.14-2(d) 结合说明, Banach 空间是自反的充分必要条件是其对偶也是自反的.

5.14-2 设 Y 是自反 Banach 空间 X 的一个闭子空间. 证明商空间 X/Y 是自反的.

5.14-3 (1) 设 X 是自反 Banach 空间. 证明, 给定任意的 $x' \in X'$, 存在 $x_0 \in X$ 使得 $\|x_0\| = 1$ 及 $\|x'\| = \sup_{\|x\|=1} |x'(x)| = x'(x_0)$.

(2) 反之, 证明如果一个 Banach 空间 X 使得对于任意的 $x' \in X'$, 都存在 $x_0 \in X$ 满足 $\|x_0\| = 1$ 而且 $\|x'\| = \sup_{\|x\|=1} |x'(x)| = x'(x_0)$, 则 X 是自反的[41)].

5.14-4 设 $(X, \|\cdot\|)$ 是自反的 Banach 空间, Z 是 X 的一个非空闭凸子集.

(1) 证明, 给定任一元素 $x \in X$, 存在 $y \in Z$ 使得 $\|x - y\| = \inf_{z\in Z} \|x - z\|$.

提示: 考虑极小化序列并利用 Banach-Eberlein-Šmulian 定理.

(2) 如果 $(X, \|\cdot\|)$ 是严格凸的, 则 y 是唯一的.

注 问题 (1) 和 (2) 扩展了 Hilbert 空间中的投影定理 (定理 4.3-1), 一起构成了自反 Banach 空间中的投影定理.

5.14-5 设装备 sup 范数 $\|\cdot\|$ 的空间 $\mathcal{C}[0,1]$ 的子集 Z 定义如下:

$$Z = \left\{ f \in \mathcal{C}[0,1]; \int_0^{1/2} f(x)dx = 1 + \int_{1/2}^1 f(x)dx \right\}.$$

(1) 证明 Z 是 $\mathcal{C}[0,1]$ 的非空闭凸子集.

(2) 证明 $\inf_{f\in Z} \|f\| = 1$, 但不存在 $f \in Z$ 使得 $\|f\| = 1$.

(3) 从习题 5.14-4 可得结论, Banach 空间 $(\mathcal{C}[0,1], \|\cdot\|)$ 不是自反的.

5.14-6 设 X 是可分的 Banach 空间. 证明 X' 中的任何有界序列都包含一个弱 * 收敛的子列.

注 这个结果构成定理 5.14-4(a) 的 "弱 * 类似".

5.14-7 设 Ω 是 $\mathbb{R}^n$ 的一个开子集, $(f_k)_{k=0}^{\infty}$ 是 $L^{\infty}(\Omega)$ 中的有界序列. 证明存在一个子列 $(f_{\sigma(k)})_{k=0}^{\infty}$ 及函数 $f \in L^{\infty}(\Omega)$ 使得

$$\text{对每个 } g \in L^1(\Omega), \quad \int_\Omega f_{\sigma(k)} g dx \to \int_\Omega f g dx, \quad \text{当 } k \to \infty \text{ 时}.$$

提示: 用习题 5.14-6.

5.14-8 设 X 是赋范向量空间. 证明, X 在标准等距 J 下的像 $J(X)$ 在 X'' 中是闭的充分必要条件为, X 是 Banach 空间.

---

41) R. C. James [1964]: Characterizations of reflexivity. Studia Mathematica **23**, 205–216.

5.14-9 设 Ω 是 $\mathbb{R}^n$ 的开子集, $1 < p < \infty$, 而函数 $f_k \in L^p(\Omega)$, $k \geqslant 1$, 及 $f \in L^p(\Omega)$ 使得序列 $(f_k)_{k=1}^{\infty}$ 在 $L^p(\Omega)$ 中有界而且当 $k \to \infty$ 时 f_k 在 Ω 中几乎处处收敛于 f. 证明

$$f_k \rightharpoonup f \quad \text{在 } L^p(\Omega) \text{ 中, 当 } k \to \infty \text{ 时.}$$

第 6 章 线性偏微分方程

引言

在这一章中, 我们只讨论所有的变量都是"空间变量"的偏微分方程, 即点的坐标均在 $\mathbb{R}^N$ 的一个开子集内. 不考虑"依赖时间"的问题.

在最优化理论中, 特别是当应用于线性化弹性或线性化流体力学时, 这类问题通常被化为下述形式的模型极小化问题: 未知量 u 满足

$$u \in U \quad 且 \quad J(u) = \inf_{v \in U} J(v),$$

其中 U 是 Hilbert 空间 V 的一个非空闭凸子集, 而 $J: V \to \mathbb{R}$ 是一个二次泛函, 即具有如下形式:

$$J(v) = \frac{1}{2} a(v, v) - \ell(v) \quad 对任意\ v \in V,$$

其中 $a(\cdot, \cdot)$ 是对称双线性形式, ℓ 是定义在空间 V 上的连续线性形式.

我们首先证明, 作为 Hilbert 空间投影定理 (第 4 章) 的一个简单的推论, 对这种极小化问题, 可以给出一个一般的存在性结果 (定理 6.1-1), 其所需的主要假定是空间 V 的完备性及双线性形式的 V 强制性. 对于这个问题, 我们还要给出另一种等价形式 (定理 6.1-2), 称其为变分形式. 它一般表示为变分不等式, 而当 U 为一子空间时, 则表示为变分方程. 当双线性形式不是对称的时, 我们单独将它们表示为抽象变分问题以讨论, 而当 $U = V$ 时, 给出一个存在性定理 (定理 6.2-1), 这就是著名的 Lax–Milgram 引理.

作为空间 V 的候选者, 应满足下述条件: 一方面它必须是完备的, 而另一方面它必须使表达式 $J(v)$ 对所有函数 $v \in V$ 是适定的. 对于我们心目中的应用问题 (弹性和流体力学) 而言, Sobolev 空间 $H^m(\Omega)$ 和 $H_0^m(\Omega)$ 满足这些要求. 这些空间, 以及 (为

了阐释的连贯性) 更一般的 Sobolev 空间 $W^{m,p}(\Omega)$ 和 $W_0^{m,p}(\Omega), 1 \leqslant p \leqslant \infty$, 的基本性质将在 6.5 和 6.6 节中做简单的回顾, 后者在最后一章相应于非二次泛函极小化的非线性偏微分方程分析中要用到. 在这两节之前, 先简单介绍一下 (6.3 节) 与分布论 (Sobolev 空间的元素本身都是分布) 有关的内容, 其中包括关于 Laplace 算子的次椭圆性的 Weyl 引理 (定理 6.4-2) 的详尽证明, 这一重要性质在以后许多地方都要用到.

在第 6.7 节到 6.9 节中, 我们将叙述并分析线性弹性中的变分问题实例, 如膜问题、板问题以及障碍问题等, 它们都符合上述抽象框架. 对于每一个例子, 主要的一步是确立有关双线性形式的 V 强制性. 作为紧自伴算子谱定理 (第 4 章) 的一个应用, 我们也将给出二阶椭圆算子特征值问题的详尽讨论 (定理 6.10-2).

我们也将阐明, 在解这种变分问题时, 是在分布意义下求解, 但如果其解具有特定的正则性, 则也是在经典意义下解二阶或四阶边值问题. 如果 U 是 V 的一个子空间, 这些问题是线性的; 而如果 U 不是 V 的一个子空间, 这些问题就是非线性的.

同样的讨论还将用于在应用上最为重要的两个线性偏微分方程组: Stokes 方程组 (6.14 节) 和三维线性化弹性方程组 (6.16 节); 顺便指出, 相应的非线性方程组 (上两个线性方程组是它们的近似) 的讨论将在第 9 章中给出. 确定这些方程组解的存在性要求更精细的分析, 它们分别依赖 Babuška-Brezzi 上下确界定理 (定理 6.12-1) 和 Korn 不等式 (定理 6.15-1), 而这两者最终都依赖同一个基本而且深刻的 Jacques–Louis Lions 引理 (定理 6.11-4).

我们以关于几个或许不太传统课题的分析讨论结束本章, 这其中包括如 Poincaré 引理, 其经典形式 (定理 6.17-2) 及其弱形式 (即在具有负指数的 Sobolev 空间中, 见定理 6.17-4); 经典和弱 Saint-Venant 引理 (定理 6.18-1 和 6.18-3); Cesàro-Volterra 路径积分公式 (定理 6.18-2); Donati 引理 (定理 6.19-5 和 6.19-6); 最后还有 Pfaff 方程组的存在唯一性定理 (定理 6.20-1), 它对于确立微分几何中的存在定理将起着关键作用 (第 8 章).

本章中所涉及的所有函数和向量空间均为实的.

6.1 二次极小化问题; 变分方程和变分不等式

我们从一个极小化问题的存在和唯一性结果开始我们的讨论, 这个问题是各种各样课题的一个模型, 本章中将对其进行充分的讨论.

要注意, 今后我们将用诸如 $u, v \in V$ 等符号, 而不用 $x, y \in X$ 等. 这类符号常用于所讨论的空间是函数空间的情况, 即空间的元素是 (如定义在 $\mathbb{R}^N$ 的一个开子集上的) 函数本身. 在下面要予以讨论的例子中就是如此. 还是根据习惯用语, 如在定理 6.1-1 中出现的函数 $J: V \to \mathbb{R}$ 下面将称为泛函.

定理 6.1-1 设 $(V, \|\cdot\|)$ 是 Banach 空间, $a: V \times V \to \mathbb{R}$ 是一个对称的连续双线

性形式. 它具有以下性质: 存在常数 α 使得

$$\alpha > 0 \quad 且 \quad a(v,v) \geqslant \alpha\|v\|^2 \quad 对所有\ v \in V.$$

设 $\ell : V \to \mathbb{R}$ 是连续线性形式, 而泛函 $J : V \to \mathbb{R}$ 由下式定义:

$$J(v) := \frac{1}{2}a(v,v) - \ell(v) \quad 对所有\ v \in V.$$

最后, 设 U 是 V 的一个非空闭凸子集.

则存在唯一的元素 u 使得

$$u \in U \quad 且 \quad J(u) = \inf_{v \in U} J(v).$$

由此种方式定义的映射 $\ell \in V' \to u \in U$ 是 Lipschitz 连续的, 当且仅当 U 是 V 的子空间时, 它是线性的.

证明　由于双线性形式 a 是连续的, 存在 $M > 0$ 使得 $|a(u,v)| \leqslant M\|u\|\|v\|$ 对所有 $u, v \in V$ 成立 (定理 2.11-1). 对称的双线性形式 $a(\cdot,\cdot)$ 显然是空间 V 上的一个内积, 且其相关的范数等价于给定的范数 $\|\cdot\|$, 此因

$$\sqrt{\alpha}\|v\| \leqslant \sqrt{a(v,v)} \leqslant \sqrt{M}\|v\| \quad 对所有\ v \in V.$$

所以, V 中装配以这个内积后, 成为 Hilbert 空间.

由 F. Riesz 表示定理 (定理 4.6-1), 存在唯一的元素 $c = c(\ell) \in V$ 使得

$$\ell(v) = a(c,v) \quad 对所有\ v \in V.$$

再考虑到双线性形式的对称性, 我们可以将泛函 J 重新表示为

$$J(v) = \frac{1}{2}a(v,v) - a(c,v) = \frac{1}{2}a(v-c, v-c) - \frac{1}{2}a(c,c).$$

所以, 找 $u \in U$ 使得 $J(u) = \inf_{v\in U} J(v)$ 就化为在范数 $\sqrt{a(\cdot,\cdot)}$ 下, 使元素 $c \in V$ 与集合 U 之间的距离达到极小的问题. 换言之, 解 u 是 c 在集合 U 上关于内积 $a(\cdot,\cdot)$ 的投影. 由投影定理 (定理 4.3-1), 又因为 U 是空间 V 的非空闭凸子集, 这个投影存在且是唯一的.

由于映射 $\ell \in V' \to c \in V$ 及 $c \in V \to u \in U$ 都是 Lipschitz 连续的 (定理 4.6-1 及 4.3-1), 其复合映射 $\ell \in V' \to u \in U$ 也是 Lipschitz 连续的.

因为映射 $\ell \in V' \to c \in V$ 是线性的, 故当且仅当投影算子 $c \in V \to u \in U$ 本身是线性的, 即当且仅当 U 是一个子空间时 (定理 4.3-1), 映射 $\ell \in V' \to u \in U$ 才是线性的. 所以, 如果 U 是一个子空间, 则映射 $\ell \in V' \to u \in U$ 是线性的; 如果 U 不是一个子空间, 则上映射为非线性的 (这里所有其他的条件均保持不变). □

注 然而, 读者不应忘记, 当在一个子空间上极小化泛函 J 时, 如果所导致的问题是线性的, 这也是因为 J 是二次的. 一个非二次的泛函在一个子空间上的极小化将导致一个非线性问题; 关于这种例子可参阅第 9 章.

设 V 是一赋范向量空间, 其中范数为 $\|\cdot\|$. 一个双线性形式 $a: V\times V\to\mathbb{R}$ (对称与否均可) 如果具有以下性质: 存在常数 α 使得

$$\alpha>0 \quad 且 \quad a(v,v)\geqslant\alpha\|v\|^2 \quad 对所有\ v\in V,$$

则称其为 V 强制的.

注 V 强制的双线性形式也经常称为 V 椭圆的. 对此应该谨慎对待, 这是因为这个概念与一致椭圆型偏微分算子和二阶椭圆边值问题 (后面将要介绍, 见 6.7 节) 中的相关概念有关, 但并不等价.

泛函 $J: V\to\mathbb{R}$ 称为二次的, 如果它具有以下形式: $J(v)=\dfrac{1}{2}a(v,v)-\ell(v)$ 对所有的 $v\in V$ 成立, 其中 $a: V\times V\to\mathbb{R}$ 是连续的对称双线性形式, 而 $\ell: V\to\mathbb{R}$ 是连续线性形式.

注 如果双线性形式是 V 强制的, 则相应的二次泛函在下述意义下也是强制的: $\lim_{\|v\|\to\infty}J(v)=\infty$ (此因 $J(v)\geqslant\dfrac{\alpha}{2}\|v\|^2-\|\ell\|\|v\|$ 对所有 $v\in V$ 成立). 更一般的强制泛函将在第 9 章中讨论.

二次极小化问题在于探寻是否存在一个元素使得二次泛函 $J: V\to\mathbb{R}$ 在 V 的一个非空子集 U 上 (的限制) 达到极小.

下面我们证明, 在定理 6.1-1 中讨论的二次极小问题可以给出另外一种等价的形式.

定理 6.1-2 假设及符号与定理 6.1-1 中相同. 一个元素 $u\in U$ 是定理 6.1-1 中的极小化问题的解之充分必要条件是它满足如下关系: 在一般的情况为

$$a(u,v-u)\geqslant\ell(v-u) \quad 对所有\ v\in U,$$

若 U 是 V 的闭子空间则为

$$a(u,v)=\ell(v) \quad 对所有\ v\in U.$$

证明 设 $c\in V$ 使得 $\ell(v)=a(c,v)$ 对所有 $v\in V$ 成立. 投影定理说明 (定理 4.3-1) $u\in U$ 是 c 到 U 上的投影 (见定理 6.1-1 的证明) 的充分必要条件是关系式

$$a(u-c,v-u)\geqslant 0 \quad 对所有\ v\in U\ 成立.$$

因为这一关系式可以重写为

$$a(u,v-u)\geqslant a(c,v-u)=\ell(v-u) \quad 对所有\ v\in U,$$

故所示的不等式成立.

如果 U 是 V 的一个子空间, 投影定理断定 $u \in U$ 是 c 到 U 上的投影的充分必要条件是

$$a(u-c,v)=0 \quad \text{对所有 } v \in U,$$

即 $a(u,v)=\ell(v)$ 对所有 $v \in U$ 成立. □

在此还要特别阐明的是, 在定理 6.1-2 中, 是通过双线性形式 a 与线性形式 ℓ 来刻画泛函 J 本身的特性. 为此, 我们利用恒等式

$$\begin{aligned} J(u+w) &= \frac{1}{2}a(u+w,u+w)-\ell(u+w) \\ &= J(u)+\{a(u,w)-\ell(w)\}+\frac{1}{2}a(w,w). \end{aligned}$$

利用 a 的双线性和对称性, 在此这是实质性的, 以及 ℓ 的线性性质, 容易验证上式对任意的 $u, w \in V$ 成立. 这个恒等式说明, 对于固定的 $u \in V$, 实数 $\{a(u,w)-\ell(w)\}$ 正是差 $J(u+w)-J(u)$ 关于 w 的线性部分. 这个线性部分称为泛函 J 在 u 的第一变分.

这样, 假定一元素 $u \in U$ 像在定理 6.1-2 中那样, 对所有的 $v \in U$ 满足 $a(u,v-u) \geqslant \ell(v-u)$. 在上述恒等式中令 $w=v-u$, 则有

$$\begin{aligned} J(v)-J(u) &= \{a(u,w)-\ell(w)\}+\frac{1}{2}a(w,w) \\ &\geqslant \frac{\alpha}{2}\|w\|^2 \quad \text{对所有 } v \in U. \end{aligned}$$

因此, $J(u)=\inf_{v \in U} J(v)$. 反之, 假定 $J(u)=\inf_{v \in U} J(v)$. 给定任意元素 $v=u+w \in U$, 我们就有 $J(u+\theta w)-J(u) \geqslant 0$ 对所有 $0 \leqslant \theta \leqslant 1$ 成立 (注意, 集合 U 是凸的). 这样就有

$$\theta\{a(u,w)-\ell(w)\}+\frac{\theta^2}{2}a(w,w) \geqslant 0 \quad \text{对所有 } 0 \leqslant \theta \leqslant 1,$$

此式意味着 $a(u,w)-\ell(w)=a(u,v-u)-\ell(v-u) \geqslant 0$ 对所有 $v \in U$ 成立.

换言之, 一个元素 $u \in U$ 使得 $J(u)=\inf_{v \in U} J(v)$ 的充分必要条件是, 对所有使得 $(u+w) \in U$ 的 $w \in V$, 泛函 J 在 u 处的第一变分 $\{a(u,w)-\ell(w)\}$ 非负.

在 U 是一个子空间的特殊情况下, 令一个元素 $u \in U$ 对所有 $w \in U$ 满足 $a(u,w)=\ell(w)$. 则上述恒等式给出

$$J(u+w)-J(u)=\frac{1}{2}a(w,w) \geqslant \frac{\alpha^2}{2}\|w\|^2 \quad \text{对所有 } w \in U.$$

所以 $J(u)=\inf_{v \in U} J(v)$. 反之, 假定 $J(u)=\inf_{v \in U} J(v)$. 给定任意元素 $v=u+w \in U$, 我们有 $J(u+\theta w)-J(u) \geqslant 0$ 对所有 $\theta \in \mathbb{R}$ 成立. 所以

$$\theta\{a(u,w)-\ell(w)\}+\frac{\theta^2}{2}a(w,w) \geqslant 0 \quad \text{对所有 } \theta \in \mathbb{R},$$

这意味着, 对所有 $w \in U$ 有 $a(u,w)-\ell(w)=0$.

换言之, 如果 U 是子空间, 一个元素 $u \in U$ 使得 $J(u) = \inf_{v \in U} J(v)$ 的充分必要条件是泛函 J 在 u 的第一变分 $\{a(u, w) - \ell(w)\}$ 对所有的 $w \in U$ 均为零.

注 在第 7 章中, 每个第一变分都将被置于其更恰当的描述方式中, 即将其视为泛函 J 在 u 处沿方向 w 的 Gâteaux 导数 $J'(u)w$.

上述讨论说明了为什么定理 6.1-2 中给出的表示称为定理 6.1-1 中的极小化问题的变分形式. 关系式 $a(u, v - u) \geqslant \ell(v - u)$ 对所有 $v \in U$ 成立, 称为变分不等式, 而关系式 $a(u, v) = \ell(v)$ 对所有 $v \in U$ 成立, 则称为变分方程.

注 在 6.12 节中, 将讨论另一类二次极小问题, 其中 U 是 V 中形式如 $U = \{v \in V; b(v, \mu) = 0$, 对所有 $\mu \in M\}$ 的闭子空间, 这里 V 与 M 都是 Hilbert 空间, $b : V \times M \to \mathbb{R}$ 是满足 Babuška-Brezzi 上下确界条件的双线性形式, 而双线性形式 $a : V \times V \to \mathbb{R}$ 只是 U 椭圆的; 见定理 6.12-2.

习题

6.1-1 设 $(V, \|\cdot\|)$ 是 Hilbert 空间, 而 $a : V \times V \to \mathbb{R}$ 是具有下述性质的对称连续的双线性形式: $a(v, v) \geqslant 0$ 对所有 $v \in V$ 成立, 而且给定任意的连续线性形式 $\ell : V \to \mathbb{R}$, 变分方程 $a(u, v) = \ell(v)$ 对所有 $v \in V$ 成立, 有且只有一个解 $u \in V$.

证明存在常数 $\alpha > 0$ 使得 $a(v, v) \geqslant \alpha \|v\|^2$ 对所有 $v \in V$ 成立.

注 这个结果构成定理 6.1-2 当 $U = V$ 时之逆.

6.2 Lax–Milgram 引理

给定向量空间 V 的一个非空子集 U, 一个双线性形式 $a(\cdot, \cdot) : V \times V \to \mathbb{R}$ 及一个线性形式 ℓ, 我们可以讨论下述抽象变分问题. 在这种形式里, 无泛函出现在其中: 在一般情况, 求一个元素 $u \in U$ 使得

$$a(u, v - u) \geqslant \ell(v - u) \quad \text{对所有 } v \in U;$$

如果 U 是一个子空间, 则求一个元素 $u \in U$ 使得

$$a(u, v) = \ell(v) \quad \text{对所有 } v \in V.$$

根据定理 6.1-2, 如果空间 V 是完备的, V 的非空子集 U 是闭凸的, 线性形式 ℓ 是连续的, 而双线性形式是 V 强制、连续并且对称的, 这些问题中的每一个均有且只有一个解.

然而可以证明, 如果把双线性形式的对称性这一假设去掉, 而 V 是 Hilbert 空间, 这种抽象变分问题仍旧有且仅有一个解. 这里, 我们只局限于 $U = V$ 这一特殊情况, 将一般情况, 构成 Stampacchia 定理, 作为习题留给读者 (习题 6.2-1).

定理 6.2-1 (Lax-Milgram 引理[1]) 设 V 是 Hilbert 空间, $a(\cdot,\cdot): V\times V\to\mathbb{R}$ 是连续且 V 强制的双线性形式, 而 $\ell: V\to\mathbb{R}$ 是连续线性形式.

抽象变分问题: 求一元素 $u\in V$ 使得

$$a(u,v)=\ell(v)\quad \text{对所有 } v\in V,$$

有且只有一个解, 且由此定义的映射 $\ell\in V'\to u\in V$ 是线性且连续的.

证明 设 $(\cdot,\cdot)$ 与 $\|\cdot\|$ 表示空间 V 中的内积与范数, 而 M 为常数使得

$$|a(u,v)|\leqslant M\|u\|\|v\|\quad \text{对所有 } u,v\in V.$$

这个关系式说明, 对于每一个 $u\in V$, 线性形式 $v\in V\to a(u,v)\in\mathbb{R}$ 是连续的. 因此, 对每个 $u\in V$, 存在唯一的元素 $Au\in V'$ 使得

$$a(u,v)=Au(v)\quad \text{对所有 } v\in V.$$

因为 $a(\cdot,\cdot)$ 关于其第一个自变量是线性的, 故如此定义的映射 $A: V\to V'$ 是线性的. 又因为

$$\|Au\|_{V'}=\sup_{v\neq 0}\frac{|Au(v)|}{\|v\|}=\sup_{v\neq 0}\frac{|a(u,v)|}{\|v\|}\leqslant M\|u\|\quad \text{对所有 } u\in V,$$

故它还是连续的. 所以

$$\|A\|_{\mathcal{L}(V;V')}\leqslant M.$$

这样, 解抽象变分问题就等价于解方程

$$Au=\ell \text{ 在 } V' \text{ 中}\quad \text{或等价地}\quad \tau(Au-\ell)=0 \text{ 在 } V \text{ 中},$$

其中 $\tau: V'\to V$ 表示 F. Riesz 映射 (定理 4.6-1).

我们现在证明, 对每个 $\ell\in V'$, 这个方程有且只有一个解 $u\in V$. 为此先证明, 对适当的值 $\rho>0$, 仿射映射

$$f_\rho: v\in V\to v-\rho\tau(Av-\ell)\in V$$

是压缩的. 设 $\alpha>0$ 使得 $a(v,v)\geqslant\alpha\|v\|^2$ 对所有 $v\in V$ 成立. 则对任意 $\rho>0$,

$$\begin{aligned}\|v-\rho\tau Av\|^2&=\|v\|^2-2\rho(\tau Av,v)+\rho^2\|\tau Av\|^2\\&\leqslant(1-2\rho\alpha+\rho^2M)\|v\|^2,\end{aligned}$$

[1] P. D. Lax; A. M. Milgram [1954]: Parabolic equations. in *Contributions to the Theory of Partial Differential Equations*, Annals of Mathematics Studies, No. 33, pp. 167–190, Princeton University Press, Princeton, NJ.

这里给出的证明取自:

J. L. Lions; G. Stampacchia [1967]: Variational inequalities. Communications on Pure and Applied Mathematics **20**, 493–519.

由于对偏微分方程理论和逼近里程碑式的贡献, Peter D. Lax 于 2005 年获得了 Abel 奖.

此因 $(\tau Av, v) = a(v, v) \geqslant \alpha\|v\|^2$ 而 $\|\tau Av\| = \|Av\|_{V'} \leqslant M\|v\|$. 所以当 ρ 属于区间 $]0, \frac{2\alpha}{M^2}[$ 时, 仿射映射 f_ρ 是压缩的. 于是由 Banach 不动点定理 (定理 3.7-1) 而得上述结论, 不动点定理说明 f_ρ 有唯一的不动点 $u \in V$, 所以满足 $\tau(Au - \ell) = 0$.

线性算子 $A \in \mathcal{L}(V; V')$ 因此是从 V 到 V' 上的双射. 逆映射 $A^{-1} : \ell \in V' \to u = A^{-1}\ell \in V$ 也是线性的 (定理 2.9-1), 并且是连续的, 此因不等式

$$\alpha\|u\|^2 \leqslant a(u, u) = \ell(u) \leqslant \|\ell\|\|u\|$$

意味着

$$\|A^{-1}\ell\| = \|u\| \leqslant \alpha^{-1}\|\ell\| \quad \text{对所有 } \ell \in V'. \qquad \square$$

注 实际上, $\|A\|_{\mathcal{L}(V;V')} = \|a\|$, 其中 $\|a\| := \sup_{\left\{\substack{u \neq 0 \\ v \neq 0}\right.} \frac{|a(u,v)|}{\|u\|\|v\|}$ 表示连续双线性形式 $a : V \times V \to \mathbb{R}$ 的范数 (定理 2.11-4). 为说明这一点, 首先注意, 由于 $|a(u, v)| \leqslant \|a\|\|u\|\|v\|$ 对所有 $u, v \in V$ 成立, 故上面的证明说明 $\|A\|_{\mathcal{L}(V;V')} \leqslant \|a\|$. 其次, 设 $u_n, v_n \in V, n \geqslant 1$, 使得对所有 $n \geqslant 1$ 有 $\|u_n\| = \|v_n\| = 1$ 且 $\|a\| = \lim_{n\to\infty} a(u_n, v_n)$; 因为对所有 $n \geqslant 1, a(u_n, v_n) = Au_n(v_n) \leqslant \|Au_n\|_{V'} \leqslant \|A\|_{\mathcal{L}(V;V')}$, 这就得到 $\|a\| \leqslant \|A\|_{\mathcal{L}(V;V')}$.

利用将于第 7 章引入的概念, 可以进一步证明, 如果双线性形式 a 不是对称的, 定理 6.2-1 中所考虑的抽象变分问题就不再有与其相应的泛函; 更准确地说, 对 $v \in V$ 表达式 $\{a(u, v) - \ell(v)\}$ 不再是一个特定泛函 J 在 u 的第一变分 (6.1 节). 这是因为, 如果这个表达式是第一变分的话, 它可以重写为 $a(u, v) - \ell(v) = J'(u)v$ 对所有 $v \in V$ 成立, 其中 $J'(u) \in V'$ 是泛函 $J : V \to \mathbb{R}$ 在 u 处的 Fréchet 导数; 而这个关系进而又意味着 J 的二阶 Fréchet 导数 $J''(u) \in \mathcal{L}_2(V; \mathbb{R})$ 由下式给出: $J''(u)(v, w) = a(v, w)$ 对所有 $(v, w) \in V \times V$ 成立. 但是可以证明 (定理 7.8-1), 任何二阶 Fréchet 导数必须对称, 这就导致矛盾.

习题

6.2-1 这个习题的问题 (1) 及 (2) 构成 Stampacchia 定理[2].

设 V 是 Hilbert 空间, $a(\cdot, \cdot) : V \times V \to \mathbb{R}$ 是连续且 V 强制的双线性形式, 又设 $\ell : V \to \mathbb{R}$ 是连续线性形式, 最后假设 U 是 V 的非空闭凸子集.

(1) 证明下列抽象变分问题: 找一元素 $u \in U$ 使得

$$a(u, v - u) \geqslant \ell(v - u) \quad \text{对所有 } v \in U$$

有且只有一个解.

(2) 证明由此定义的映射 $\ell \in V' \to u \in V$ 是 Lipschitz 连续的.

---

[2] G. Stampacchia [1964]: Formes bilinéaires coercitives sur les ensembles convexes. Comptes Rendus de l'Académis des Sciences de Paris Série A, **258**, 4413–4416.

提示: 对于 (1), 仿效定理 6.2-1 的证明; 对于 (2), 证明如果 $u_1, u_2 \in U$ 是相应于 $\ell_1, \ell_2 \in V'$ 的解, 则 $\|u_1 - u_2\| \leqslant \alpha^{-1}\|\ell_1 - \ell_2\|$.

6.2-2 设 $(V, \|\cdot\|)$ 是 Hilbert 空间, 而 $a(\cdot,\cdot): V \times V \to \mathbb{R}$ 是具有下述性质的连续双线性形式: 给定任意连续线性形式 $\ell: V \to \mathbb{R}$ 及 V 的任意闭子空间 U, 变分方程 $a(u, v) = \ell(v)$ 对所有 $v \in U$ 有且只有一个解 $u \in U$.

证明存在常数 $\alpha > 0$ 使得 $a(v, v) \geqslant \alpha\|v\|^2$ 对所有 $v \in V$ 或 $-a(v, v) \geqslant \alpha\|v\|^2$ 对所有 $v \in V$ 成立.

注 这个结果[3] 构成 Lax-Milgram 引理 (定理 6.2-1) 之逆.

6.3 $L^1_{\text{loc}}(\Omega)$ 中的弱偏导数; 分布论简介

本节的目的是介绍弱偏导数的概念, 其在 Sobolev 空间的定义 (6.5 节) 中起着决定性的作用. 我们也将说明为什么这种弱导数实际上是分布意义下导数的特殊情况.

设 Ω 是 $\mathbb{R}^N$ 的开子集. 我们记得, $\mathcal{D}(\Omega)$ 表示所有这样的无穷阶可微函数 $\varphi: \Omega \to \mathbb{R}$, 其支集 $\operatorname{supp}\varphi$ 是 Ω 的紧子集 (2.6 节) 组成的空间; $L^1_{\text{loc}}(\Omega)$ 表示所有这样的可测函数 $v: \Omega \to \mathbb{R}$, 对 Ω 中任一紧子集 K (2.6 节), 其在 K 的限制 $v|_K \in L^1(K)$ 组成的空间.

我们首先证明一个简单但非常有用的公式, 它被 $\mathbb{R}^N$ 的一个开子集上的 $\mathcal{C}^m$ 类函数的常规偏导数所满足. 这个公式 (当只涉及具有紧支集的函数时) 可视为无边界项的分部积分公式, 将为下面定义 m 阶弱偏导数提供依据.

定理 6.3-1 设 Ω 为 $\mathbb{R}^N$ 的开子集, $m \geqslant 1$ 是整数, 而函数 $u \in \mathcal{C}^m(\Omega)$ 是给定的. 则

$$\int_\Omega (\partial^{\boldsymbol{\alpha}} v)\varphi dx = (-1)^{|\boldsymbol{\alpha}|}\int_\Omega v\partial^{\boldsymbol{\alpha}}\varphi dx \quad 对所有\ \varphi \in \mathcal{D}(\Omega),$$

对每个阶数 $|\boldsymbol{\alpha}| \leqslant m$ 的重指标 $\boldsymbol{\alpha}$ 成立.

证明 设给定函数 $v \in \mathcal{C}^1(\Omega)$. 则对任意函数 $\varphi \in \mathcal{D}(\Omega)$, 每一个积分 $\int_\Omega (\partial_i v)\varphi dx$, $1 \leqslant i \leqslant N$, 都是有意义的.

函数 $w := v\varphi \in \mathcal{C}^1(\Omega)$ 在 Ω 中具有紧支集. 设 $\widehat{w}$ 表示 w 在 $\mathbb{R}^N - \Omega$ 中取零值的延拓, 这就有 $\widehat{w} \in \mathcal{C}^1(\mathbb{R}^N)$. 因为 $\operatorname{supp}\widehat{w} = \operatorname{supp} w \subset \operatorname{supp}\varphi$, 故存在 $a > 0$ 使 $\operatorname{supp}\widehat{w} \subset]-a, a[^N$. 所以, 由 Fubini 定理 (定理 1.15-5),

$$\begin{aligned}\int_\Omega \partial_i w dx &= \int_{[-a,a]^N} \partial_i \widehat{w} dx \\ &= \int_{[-a,a]^{N-1}}\left(\int_{-a}^{a} \partial_i \widehat{w}(x_1, \cdots, x_{i-1}, t, x_{i+1}, \cdots, x_N)dt\right)dx_1\cdots dx_{i-1}dx_{i+1}\cdots dx_N,\end{aligned}$$

[3] 应归于 Luc Tartar (私人交流).

因此有

$$\int_\Omega \partial_i w dx = \int_\Omega (\partial_i v)\varphi dx + \int_\Omega v\partial_i \varphi dx = 0,$$

这是由于对于 $t=-a$ 及 $t=a, \widehat{w}(x_1,\cdots,x_{i-1},t,x_{i+1},\cdots,x_N)=0$. 所以

$$\int_\Omega (\partial_i v)\varphi dx = -\int_\Omega v\partial_i\varphi dx \quad 对所有\ \varphi\in\mathcal{D}(\Omega),$$

这就在 $|\boldsymbol{\alpha}|=1$ 的情况证明了定理.

对阶数 $|\boldsymbol{\alpha}|\geqslant 2$ 的任意偏微分算子 $\partial^{\boldsymbol{\alpha}}$, 可类似地予以证明. □

受定理 6.3-1 中确立的公式的启发, 给出下面的基本定义: 给定函数 $v\in L^1_{\text{loc}}(\Omega)$ 及任意重指标 $\boldsymbol{\alpha}$, $|\boldsymbol{\alpha}|\geqslant 1$, 一个函数 $v^{\boldsymbol{\alpha}}\in L^1_{\text{loc}}(\Omega)$ 如果满足

$$\int_\Omega v^{\boldsymbol{\alpha}}\varphi dx = (-1)^{|\boldsymbol{\alpha}|}\int_\Omega v\partial^{\boldsymbol{\alpha}}\varphi dx \quad 对所有\ \varphi\in\mathcal{D}(\Omega),$$

则称其为 v 在 $L^1_{\text{loc}}(\Omega)$ 中的 $|\boldsymbol{\alpha}|$ 阶弱偏导数. 例如, 给定一个函数 $v\in L^1_{\text{loc}}(\Omega)$, 一个函数 $v_i\in L^1_{\text{loc}}(\Omega)$ 是 v 的在 $L^1_{\text{loc}}(\Omega)$ 中关于第 i 个变量的一阶弱偏导数, 如果

$$\int_\Omega v_i\varphi dx = -\int_\Omega v\partial_i\varphi dx \quad 对所有\ \varphi\in\mathcal{D}(\Omega).$$

我们在定理 6.3-3 中将会看到, 上述弱偏导数定义的合理性有赖于下述关于空间 $L^1_{\text{loc}}(\Omega)$ 中函数的 (就其本身也是很重要的) 性质. 这一性质在变分学中也起着基本作用 (9.1 节), 下面定理的名字也反映了这一点.

定理 6.3-2 (变分学的基本引理) 设 Ω 是 $\mathbb{R}^N$ 的开子集. 若一函数 $v\in L^1_{loc}(\Omega)$ 使得

$$\int_\Omega v\varphi dx = 0 \quad 对所有\ \varphi\in\mathcal{D}(\Omega),$$

则 $v=0$.

证明 定义开集

$$\Omega_k := \left\{x\in\Omega; \operatorname{dist}(x,\mathbb{R}^N-\Omega) > \frac{1}{k}\right\}\cap B(0;k) \quad 对每一整数\ k\geqslant 1.$$

则 $\Omega=\cup_{k=1}^\infty\Omega_k$ 而且对每个 $k\geqslant 1, \overline{\Omega}_k$ 是 Ω 的紧子集. $v\in L^1_{\text{loc}}(\Omega)$ 这一假设意味着 $v|_{\Omega_k}\in L^1(\Omega_k)$ 对每个 $k\geqslant 1$ 成立.

对每个整数 $k\geqslant 1$, 设函数 $v^k\in L^1(\Omega)$ 由下式定义:

$$v^k := v|_{\Omega_{2k}}\ 在\ \Omega_{2k}\ 上,\ v^k := 0\ 在\ \Omega-\Omega_{2k}\ 上,$$

又设 $\varepsilon_0(k)>0$ 使得

$$\overline{\Omega}_{2k}\subset\Omega_\varepsilon := \{x\in\Omega; \operatorname{dist}(x,\mathbb{R}^N-\Omega)>\varepsilon\} \quad 对所有\ 0<\varepsilon\leqslant\varepsilon_0,$$

而 $(v^k_\varepsilon)_{\varepsilon>0}$ 表示函数 v^k 的正则化族 (2.6 节). 则由定理 2.6-4 得

$$\|v^k_\varepsilon - v^k\|_{L^1(\Omega_k)} \to 0 \quad \text{当 } \varepsilon \to 0 \text{ 时}.$$

因为若 $x \in \Omega_k$, 用以定义正则化族的光滑化子 $\omega_\varepsilon : y \in \Omega \to \omega_\varepsilon(x-y)$ 属于空间 $\mathcal{D}(\Omega)$, 对函数 v 的假设意味着, 如果 $\varepsilon > 0$ 充分小, 有

$$\begin{aligned} v^k_\varepsilon(x) &= \int_{B(x;\varepsilon)} v^k(y)\omega_\varepsilon(x-y)dy \\ &= \int_\Omega v(y)\omega_\varepsilon(x-y)dy = 0 \quad \text{对每个 } x \in \Omega_k. \end{aligned}$$

所以 $v|_{\Omega_k} = 0$, 此因

$$\|v\|_{L^1(\Omega_k)} = \lim_{\varepsilon\to 0} \|v^k_\varepsilon\|_{L^1(\Omega_k)} = 0.$$

关系式 $\Omega = \cup_{k=1}^\infty \Omega_k$ 和关系式 $v|_{\Omega_k} = 0, k \geqslant 1$, 意味着 $v = 0$ 在 Ω 内 (要说明这一点, 注意由 Fatou 引理 $\int_\Omega |v|dx \leqslant \liminf_{k\to\infty} \int_{\Omega_k} |v|dx = 0$). □

注 在 $v \in L^p(\Omega), 1 < p < \infty$, 的特殊情况, 定理 6.3-2 可由空间 $\mathcal{D}(\Omega)$ 在空间 $L^q(\Omega)$ 中的稠密性 (定理 2.6-2) 和空间 $L^q(\Omega)$ 中的 F. Riesz 表示定理 (定理 3.5-3) 得到, 其中 q 表示 p 的共轭指数.

借助于定理 6.3-2, 我们现在可以来证明关于弱偏导数的两条预期应具有的性质, 即其定义是明确无歧义的而且实际上是常规偏导数概念的推广.

定理 6.3-3 设 Ω 是 $\mathbb{R}^N$ 的开子集. 给定函数 $v \in L^1_{\text{loc}}(\Omega)$ 及重指标 $\boldsymbol{\alpha}, |\boldsymbol{\alpha}| \geqslant 1$, 设函数 $v^{\boldsymbol{\alpha}} \in L^1_{\text{loc}}(\Omega)$ 是 v 的 $|\boldsymbol{\alpha}|$ 阶弱偏导数, 即满足

$$\int_\Omega v^{\boldsymbol{\alpha}}\varphi dx = (-1)^{|\boldsymbol{\alpha}|} \int_\Omega v\partial^{\boldsymbol{\alpha}}\varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega).$$

则这个弱偏导数是唯一的, 而如果 $v \in \mathcal{C}^{|\boldsymbol{\alpha}|}(\Omega)$, 则 $v^{\boldsymbol{\alpha}} = \partial^{\boldsymbol{\alpha}} v$.

证明 设 $v^{\boldsymbol{\alpha}} \in L^1_{\text{loc}}(\Omega)$ 及 $w^{\boldsymbol{\alpha}} \in L^1_{\text{loc}}(\Omega)$ 使得

$$\int_\Omega v^{\boldsymbol{\alpha}}\varphi dx = (-1)^{|\boldsymbol{\alpha}|} \int_\Omega v\partial^{\boldsymbol{\alpha}}\varphi dx = \int_\Omega w^{\boldsymbol{\alpha}}\varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega).$$

则由定理 6.3-2, 有 $v^{\boldsymbol{\alpha}} = w^{\boldsymbol{\alpha}}$.

因为如果 $v \in \mathcal{C}^{|\boldsymbol{\alpha}|}(\Omega)$, 则有

$$\int_\Omega (\partial^{\boldsymbol{\alpha}} v)\varphi dx = (-1)^{|\boldsymbol{\alpha}|} \int_\Omega v\partial^{\boldsymbol{\alpha}}\varphi dx = \int_\Omega v^{\boldsymbol{\alpha}}\varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega)$$

(定理 6.3-1). 在这种情况, 仍用定理 6.3-2 就得到 $v^{\boldsymbol{\alpha}} = \partial^{\boldsymbol{\alpha}} v$. □

我们现在证明空间 $L^1_{\text{loc}}(\Omega)$ 中函数的一条重要性质. 这一性质推广了关于连续可微函数的一条众所周知的性质, 即一个函数 $v \in \mathcal{C}^1(\Omega)$ 使得 $\partial_i v = 0, 1 \leqslant i \leqslant N$, 在 $\mathbb{R}^N$ 的连通开子集 Ω 中成立, 则其必为常数. 实际上, $\int_\Omega v \partial_i \varphi dx = 0$ 对所有 $\varphi \in \mathcal{D}(\Omega), 1 \leqslant i \leqslant N$, 成立这个假设就意味着 v 的所有一阶弱偏导数在 $L^1_{\text{loc}}(\Omega)$ 中为零 (其推广, 见习题 6.3-1).

$L^1_{\text{loc}}(\Omega)$ 中函数的另一条重要性质将在下一节给出 (定理 6.4-2).

定理 6.3-4 设 Ω 为 $\mathbb{R}^N$ 中的连通开子集, 而函数 $v \in L^1_{\text{loc}}(\Omega)$ 使得

$$\int_\Omega v \partial_i \varphi dx = 0 \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega), \quad 1 \leqslant i \leqslant N.$$

则 v 是一个常函数.

证明 由定理 1.9-2, 我们只需证明函数在 Ω 中是局部常数 (即给定任一点 $x \in \Omega$, 存在 x 的一个邻域使 v 在其中为常数), 此因 Ω 是连通的.

给定任意点 $x \in \Omega$, 存在 $r > 0$ 使得 $\overline{U} \subset \Omega$, 其中 $U := B(x; r)$. 然后设 $(v_\varepsilon)_{\varepsilon>0}$ 是给定函数 $v \in L^1_{\text{loc}}(\Omega)$ 的正则化族 (2.6 节). 因为 $\overline{U}$ 是 Ω 的紧子集, 定理 2.6-1 及 2.6-4 说明, 存在 $\varepsilon_1 = \varepsilon_1(U) > 0$ 使得对所有 $0 < \varepsilon \leqslant \varepsilon_1$,

$$\overline{U} \subset \Omega_\varepsilon := \{x \in \Omega; \text{dist}(x, \mathbb{R}^N - \Omega) > \varepsilon\}, \quad v_\varepsilon \in \mathcal{D}(\Omega_\varepsilon), \quad \|v_\varepsilon - v\|_{L^1(U)} \underset{\varepsilon \to 0}{\longrightarrow} 0,$$

$$\partial_i v_\varepsilon(x) = \int_\Omega \partial_i \omega_\varepsilon(x - y) v(y) dy \quad \text{对所有 } x \in \Omega_\varepsilon, 1 \leqslant i \leqslant N.$$

因为, 对每个 $x \in \Omega_\varepsilon$, 每一函数 $y \in \Omega \to \partial_i \omega_\varepsilon(x - y), 1 \leqslant i \leqslant N$, 都属于空间 $\mathcal{D}(\Omega)$, 对函数 v 所做的假定意味着对所有 $0 < \varepsilon \leqslant \varepsilon_1$, 成立

$$\partial_i v_\varepsilon(x) = 0 \quad \text{对所有 } x \in B(x; r), \quad 1 \leqslant i \leqslant N.$$

由微积分中的经典结果, 每一个限制 $v_\varepsilon|_{B(x;r)}, 0 < \varepsilon \leqslant \varepsilon_1$, 在连通开集 $B(x; r)$ 上都是常函数. 所以, 限制 $v|_U$ 也是常函数, 这是因为常函数 $v_\varepsilon|_U$ 当 $\varepsilon \to 0$ 时在 $L^1(U)$ 中收敛于 $v|_U$ (函数 $v_\varepsilon|_U, 0 < \varepsilon \leqslant \varepsilon_1$, 属于空间 $L^1(U)$ 的一维闭子空间 $P_0(U)$, 且在 $L^1(U)$ 中收敛). □

遵循普遍的用法, 从现在起, 对于经典及弱偏导数我们将用同样的符号表示之, 即用 $\partial_i v, \partial_{ij} v$ 等或重指标用 $\partial^\alpha v$. 但特别要注意, 不要轻率地将经典偏导数的性质移到弱偏导数中来. 当然经典偏导数的某些性质可能仍然保持, 但要确定这一点, 通常是要求专门予以证明的. 定理 6.3-4 及习题 6.3-1 就提供了这种例子.

在结束本节之前, 我们 (非常简明地) 介绍一下分布的理论基础, 在偏微分方程, 或 Laplace 及 Fourier 变换的现代理论中, 随处可见其踪迹. 对其更深入的内容 (如空间 $\mathcal{D}(\Omega)$ 及 $\mathcal{D}'(\Omega)$ 的拓扑之精确定义等) 感兴趣的读者可参阅在"文献注释"中所建议的文献.

设 Ω 是 $\mathbb{R}^N$ 中的开子集, Ω 上的 Schwartz 分布[4] 是一个线性泛函 $T:\mathcal{D}(\Omega)\to\mathbb{R}$, 具有以下性质: 给定 Ω 的任一紧子集 K, 存在常数 $C(K)$ 及整数 $m(K)\geqslant 0$, 使得

$$|T(\varphi)|\leqslant C(K)\sup_{\left\{\begin{smallmatrix}|\boldsymbol{\alpha}|\leqslant m(K)\\ x\in K\end{smallmatrix}\right.}|\partial^{\boldsymbol{\alpha}}\varphi(x)|\quad \text{对所有 } \operatorname{supp}\varphi\subset K \text{ 的 } \varphi\in\mathcal{D}(\Omega).$$

Ω 上所有分布形成的空间表示为

$$\mathcal{D}'(\Omega).$$

空间 $\mathcal{D}(\Omega)$ 以自然的方式装备以 "归纳极限" 拓扑, 这使其成为一个 "局部凸拓扑向量空间". 在这种拓扑下, 函数 $\varphi_k\in\mathcal{D}(\Omega)$ 的序列 $(\varphi_k)_{k=1}^{\infty}$ 收敛于函数 $\varphi\in\mathcal{D}(\Omega)$ 的充分必要条件是存在 Ω 的紧子集 K 使得

$$\operatorname{supp}\varphi_k\subset K\quad \text{对所有 } k\geqslant 1,\quad \operatorname{supp}\varphi\subset K,$$

而且对任一重指标 $\boldsymbol{\alpha},|\boldsymbol{\alpha}|\geqslant 0$, 当 $k\to\infty$ 时成立

$$\sup_{x\in K}|\partial^{\boldsymbol{\alpha}}\varphi_k(x)-\partial^{\boldsymbol{\alpha}}\varphi(x)|\to 0.$$

然而, 要注意空间 $\mathcal{D}(\Omega)$ 的拓扑是不可距离化的, 所以更别说赋范了.

可以证明, 空间 $\mathcal{D}'(\Omega)$ (如上面定义的) 在下述意义下是 $\mathcal{D}(\Omega)$ 的对偶空间: $\mathcal{D}'(\Omega)$ 由所有关于 $\mathcal{D}(\Omega)$ 归纳极限拓扑连续的 $\mathcal{D}(\Omega)$ 上的线性泛函组成. 作为对偶空间, $\mathcal{D}'(\Omega)$ 以自然的方式装配以 "类 - 弱 *" 拓扑 (5.12 节), 它仍是不可距离化的. 在这种拓扑下, 分布 $T^k\in\mathcal{D}'(\Omega)$ 的序列 $(T^k)_{k=1}^{\infty}$ 收敛于分布 $T\in\mathcal{D}(\Omega)$ 的充分必要条件为

$$T^k(\varphi)\underset{k\to\infty}{\longrightarrow}T(\varphi)\quad \text{对所有 } \varphi\in\mathcal{D}(\Omega).$$

如果此式成立, 则称序列 (T^k) 在分布意义下收敛于 T, 此收敛记为

$$T^k\underset{k\to\infty}{\longrightarrow}T \text{ 在 } \mathcal{D}'(\Omega) \text{ 中}.$$

给定任意 $T\in\mathcal{D}'(\Omega)$ 及任意 $\varphi\in\mathcal{D}(\Omega)$, 我们也常用下述表示:

$${}_{\mathcal{D}'(\Omega)}\langle T,\varphi\rangle_{\mathcal{D}(\Omega)}:=T(\varphi)\quad \text{或简记为}\quad \langle T,\varphi\rangle:=T(\varphi).$$

给定任意函数 $v\in L^1_{\text{loc}}(\Omega)$, 线性泛函

$$T_v:\varphi\in\mathcal{D}(n)\to T_v(\varphi):=\int_{\Omega}v\varphi dx$$

---

[4] 冠名源自 Laurent Schwartz (1915—2002) 及其里程碑式著作: SCHWARTZ [1966]. 作为一位数学家 (他于 1950 年荣获 Fields 奖)、一位教授及一个极具同情心的人, 其所取得的骄人成就的生动叙述可见其自传: SCHWARTZ [2001].

定义 Ω 上一个分布, 此因对于 Ω 的任一紧子集 K 及对任一 $\mathrm{supp}\ \varphi \subset K$ 的函数 $\varphi \in \mathcal{D}(\Omega)$, 有

$$|T_v(\varphi)| \leqslant \|v\|_{L^1(K)} \sup_{x\in K} |\varphi(x)|.$$

分布 T_v 称为与局部可积函数 v 相应的分布.

然而, 确实存在不与任何局部可积函数相应的分布. 例如, 考虑线性泛函

$$\delta_a : \varphi \in \mathcal{D}(\Omega) \to \delta_a(\varphi) := \varphi(a),$$

其中 a 是 Ω 中任一点. 因为

$$|\delta_a(\varphi)| \leqslant |\sup_{x\in K} \varphi(x)|$$

对 Ω 的任一紧子集 K 及 $\mathrm{supp}\ \varphi \subset K$ 的任一函数 $\varphi \in \mathcal{D}(\Omega)$ 成立, 故 δ_a 是一个分布, 称为在 a 点的 Dirac 分布[5], 如果 $a = 0$ 就简称为 Dirac 分布.

但是不存在任何函数 $v \in L^1_{\mathrm{loc}}(\Omega)$ 使得 $\varphi(a) = \displaystyle\int_\Omega v\varphi dx$ 对所有 $\varphi \in \mathcal{D}(\Omega)$ 成立. 为说明这一点, 考虑如下定义的函数 $\varphi_k \in \mathcal{D}(\mathbb{R}^N), k \geqslant 1$:

$$\varphi_k(x) = e^{\frac{1}{|k(x-a)|^2-1}} \quad 若\ |x-a| < \frac{1}{k}$$

及

$$\varphi_k(x) = 0 \quad 若\ |x-a| \geqslant \frac{1}{k},$$

这样, 对某个整数 $k_0 \geqslant 1, \varphi_k|_\Omega \in \mathcal{D}(\Omega)$ 对所有 $k \geqslant k_0$ 成立. 记

$$U := \{x \in \Omega; \varphi_{k_0}(x) \neq 0\}.$$

则 $v \in L^1(U), |v\varphi_k| \leqslant |v|$ 对所有 $k \geqslant k_0$ 在 Ω 中几乎处处成立, 而且当 $k \to \infty$ 时 $v\varphi_k \to 0$ 在 U 中几乎处处成立. 由 Lebesgue 控制收敛定理 (定理 1.15-3), 当 $k \to \infty$ 时, $\displaystyle\int_\Omega v\varphi_k dx \to 0$, 但对所有 $k \geqslant 1, \varphi_k(a) = e^{-1}$, 导致矛盾.

分布的广泛的来源是在分布意义下求导: 设 T 为 Ω 上的一个分布, 而 $\boldsymbol{\alpha}$ 是任意重指标, $|\boldsymbol{\alpha}| \geqslant 1$. 则由下式定义的线性泛函

$$\partial^{\boldsymbol{\alpha}} T : \varphi \in \mathcal{D}(\Omega) \to \partial^{\boldsymbol{\alpha}} T(\varphi) := (-1)^{|\boldsymbol{\alpha}|} T(\partial^{\boldsymbol{\alpha}} \varphi)$$

仍是 Ω 上的一个分布 (这是分布定义的一个简单的推论), 称为 T 在分布意义下的 $|\alpha|$ 阶偏导数. 例如, $\mathbb{R}$ 上的 Dirac 分布 (上面已证明, 它不可能相应于 $\mathbb{R}$ 上的任何局部可积函数) 是如下定义的局部可积函数 $v : \mathbb{R} \to \mathbb{R}$ 在分布意义下的导数: $v(x) = 0$ 当 $x < 0$ 时; $v(x) = 1$ 当 $x \geqslant 0$ 时 (习题 6.3-3).

[5] 冠名源自杰出的物理学家 Paul Dirac (1902—1984), 他荣获 1933 年度 Nobel 物理学奖.

更一般地, 给定重指标的任意有限集 $\boldsymbol{A}$ 及系数 $a_{\boldsymbol{\alpha}} \in \mathbb{R}, \boldsymbol{\alpha} \in \boldsymbol{A}$, 线性泛函

$$\mathcal{L}T : \varphi \in \mathcal{D}(\Omega) \to \mathcal{L}T(\varphi) := \sum_{\boldsymbol{\alpha} \in \boldsymbol{A}} (-1)^{|\boldsymbol{\alpha}|} a_{\boldsymbol{\alpha}} T(\partial^{\boldsymbol{\alpha}} \varphi) \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega),$$

定义一个分布意义下的线性偏微分算子, 即

$$\mathcal{L} := \sum_{\boldsymbol{\alpha} \in \boldsymbol{A}} a_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}} : T \in \mathcal{D}'(\Omega) \to \mathcal{L}T \in \mathcal{D}'(\Omega)$$

(为了简单, 在此只考虑常系数的偏微分算子).

给定任意分布 $f \in \mathcal{D}'(\Omega)$, 我们可以讨论是否存在一个分布 $T \in \mathcal{D}'(\Omega)$, 满足

$$\mathcal{L}T = f \quad \text{在 } \mathcal{D}'(\Omega) \text{ 中}.$$

如果存在, T 就称为 $\mathcal{L}T = f$ 在分布意义下的解. 当 $\mathcal{L} = -\Delta$ 而 $f = \delta$ 时, 在习题 6.3-4 中给出这种解 T 的一个实例 (在此例中, $T = T_v, v \in \mathcal{L}^1_{\text{loc}}(\Omega)$).

习题

6.3-1 设 Ω 为 $\mathbb{R}^N$ 中的连通开子集, 而函数 $v \in L^1_{\text{loc}}(\Omega)$ 使得对某整数 $m \geqslant 1$

$$\int_\Omega v \partial^{\boldsymbol{\alpha}} \varphi dx = 0 \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega) \text{ 及所有重指标 } \boldsymbol{\alpha}, |\boldsymbol{\alpha}| = m \text{ 成立}.$$

换言之, v 的所有 m 阶弱偏导数为零. 证明 v 是 N 个变量的阶数 $\leqslant m-1$ 的多项式 ($m = 1$ 的特殊情况, 已在定理 6.3-4 中给出证明).

6.3-2 设 Ω 是 $\mathbb{R}^N$ 中的开子集而函数 $v \in L^1_{\text{loc}}(\Omega)$ 使得

$$\int_\Omega v \varphi dx = 0 \quad \text{对所有满足 } \int_\Omega \varphi dx = 0 \text{ 的 } \varphi \in \mathcal{D}(\Omega).$$

证明 v 是常函数.

6.3-3 设 I 为 $\mathbb{R}$ 的任一包含原点的开区间.

(1) 证明函数 $v : x \in I \to v(x) := \max\{0, x\}$ (其显然属于 $L^1_{\text{loc}}(I)$) 有属于 $L^1_{\text{loc}}(I)$ 的弱导数.

(2) 证明 T_v 在分布意义下的二阶导数是 Dirac 分布 (因此 v 不具有属于 $L^1_{\text{loc}}(I)$ 的弱二阶导数).

6.3-4 设 Ω 为 $\mathbb{R}^N$ 的任一包含原点的开子集.

(1) 设 ω_N 表示 $\mathbb{R}^N$ 中单位球的体积. 证明, 在 Ω 中由下式几乎处处定义的函数 $v : \Omega \to \mathbb{R}$

$$v(x) := \frac{1}{2\omega_2} \ln |x| \quad \text{若 } x \neq 0 \text{ 且 } N = 2$$

或

$$v(x) := \frac{1}{N(2-N)\omega_N} |x|^{2-N} \quad \text{若 } x \neq 0 \text{ 且 } N \geqslant 3,$$

属于空间 $L^1_{\text{loc}}(\Omega)$.

(2) 证明对任何 $N \geqslant 2$

$$\int_\Omega v \Delta\varphi dx = \varphi(0) \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega),$$

即

$$\Delta v = \delta_0 \quad \text{在分布意义下}.$$

为此原因, 函数 v 称为 Laplace 方程的基本解.

6.4 Δ 的次椭圆性

下述结论是解析函数理论中熟知的一个结果: 设 Ω 是 $\mathbb{R}^2$ 中的开子集; 则任何在 Ω 中满足 Laplace 方程 $\Delta v = 0$ 的函数 $v \in \mathcal{C}^2(\Omega)$ 实际上是解析的, 因此特别地, 在 Ω 中属于 $\mathcal{C}^\infty$ 类.

值得关注的是, 对于这一结果可给出如下多方位的推广, 它对任何维数, 在更一般的分布意义下, 以及对更一般的 Poisson 方程都成立: 设 Ω 为 $\mathbb{R}^N$ 的一个开子集, 那么任何分布 $T \in \mathcal{D}'(\Omega)$, 如果它满足

$$\Delta T = f \quad \text{在 } \mathcal{D}'(\Omega) \text{ 中, 其中 } f \in \mathcal{C}^\infty(\Omega),$$

即满足 (6.3 节)

$$T(\Delta\varphi) = \int_\Omega f\varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega),$$

实际上是一个函数, 而且也属于空间 $\mathcal{C}^\infty(\Omega)$. 这一性质称为 Δ 的次椭圆性, 要在这种一般性的水平上予以证明, 并不容易 [6) 7)].

在此, 我们对分布 T 是 Ω 中的*局部可积函数*这一 (重要的) 特殊情况给出证明; 对这种函数而言, Δ 的次椭圆性稍后将被用于证明一个 "弱" Poincaré 引理 (定理 6.17-4).

为此目的, 我们首先证明几个就其本身也是很有意义的结果. 它们在一定意义上使我们联想起前面在定理 2.6-1(b) 和 2.6-3 中证明的结果. 但要注意的是, 与前面引入的光滑化子不同, 在下述定理中考虑的函数 ρ_ε 不再要求有紧支集.

下面, 符号 B_δ 表示 $\mathbb{R}^N$ 中球心在原点, 半径为 δ 的开球; 给定 $\mathbb{R}^N$ 中的两个开子集 U 与 V, 符号 $U \subset\subset V$ 意指 $\overline{U}$ 是 V 的紧子集.

[6)] 证明例如可参见 Vo-Khac [1972b, DB 章第 3 节].

[7)] 更一般地, 对于常系数线性偏微分算子 $\mathcal{L}$, 如果每一个使得 $\mathcal{L}v \in \mathcal{C}^\infty(\Omega)$ 的函数 $v \in L^1_{\text{loc}}(\Omega)$ 本身也一定属于 $\mathcal{C}^\infty$ 类, 则称其为*次椭圆的*. 对这种算子的次椭圆性的充分必要条件在下述文献中给出:

L. Hörmander [1955]: On the theory of general differential operators. Acta Mathematica **94**, 161–248.

另一个证明, 依赖闭图像定理 (5.7 节), 可参阅 Yosida [1966, 第 2 章第 7 节].

定理 6.4-1 设 $(\rho_\varepsilon)_{\varepsilon>0}$ 是具有下述性质的函数 $\rho_\varepsilon \in L^1(\mathbb{R}^N)$ 的函数族:

$$\rho_\varepsilon(x) \geqslant 0 \quad \text{对所有 } x \in \mathbb{R}^N, \quad \int_{\mathbb{R}^N} \rho_\varepsilon(y)dy = 1,$$

$$\text{对每个 } \delta > 0, \quad \int_{\mathbb{R}^N - B_\delta} \rho_\varepsilon(y)dy \to 0 \quad \text{当 } \varepsilon \to 0 \text{ 时}.$$

(a) 设 $w : \mathbb{R}^N \to \mathbb{R}$ 是有界的一致连续函数. 则对每个 $\varepsilon > 0$, 由下式定义的函数 $w \star \rho_\varepsilon : \mathbb{R}^N \to \mathbb{R}$,

$$(w \star \rho_\varepsilon)(x) := \int_{\mathbb{R}^N} w(x-y)\rho_\varepsilon(y)dy \quad \text{对每个 } x \in \mathbb{R}^N$$

也是有界的, 且

$$\sup_{x\in\mathbb{R}^N} |(w \star \rho_\varepsilon)(x) - w(x)| \to 0 \quad \text{当 } \varepsilon \to 0 \text{ 时}.$$

(b) 给定一个函数 $v \in L^1(\mathbb{R}^N)$. 则对每个 $\varepsilon > 0, v \star \rho_\varepsilon \in L^1(\mathbb{R}^N)$. 此外, 给定任意开集 $V \subset\subset \mathbb{R}^N$,

$$\|v \star \rho_\varepsilon - v\|_{L^1(V)} \to 0 \quad \text{当 } \varepsilon \to 0 \text{ 时}.$$

证明 (i) 设 $w : \mathbb{R}^N \to \mathbb{R}$ 是满足 (a) 中假设的一个函数. 因为 $\rho_\varepsilon \geqslant 0$ 且 $\int_{\mathbb{R}^N} \rho_\varepsilon dx = 1$,

$$\begin{aligned} |(w \star \rho_\varepsilon)(x)| &\leqslant \int_{\mathbb{R}^N} |w(x-y)|\rho_\varepsilon(y)dy \\ &\leqslant \sup_{z\in\mathbb{R}^N} |w(z)| \quad \text{对每个 } x \in \mathbb{R}^N. \end{aligned}$$

故函数 $w \star \rho_\varepsilon$ 是有界的.

给定任意的 $\eta > 0$, 存在 $\delta = \delta(\eta) > 0$ 及 $\varepsilon_0 = \varepsilon_0(\delta) = \varepsilon_0(\eta) > 0$ 使得

$$|w(x-y) - w(x)| \leqslant \frac{1}{2}\eta \quad \text{对所有 } x \in \mathbb{R}^N \text{ 及所有 } y \in B_\delta,$$

$$\left(\sup_{z\in\mathbb{R}^N} |w(z)|\right) \int_{\mathbb{R}^N - B_\delta} \rho_\varepsilon(y)dy \leqslant \frac{1}{4}\eta \quad \text{对所有 } \varepsilon \leqslant \varepsilon_0.$$

因此对任何 $x \in \mathbb{R}^N$,

$$\begin{aligned} |(w \star \rho_\varepsilon)(x) - w(x)| &= \left| \int_{\mathbb{R}^N} (w(x-y) - w(x))\rho_\varepsilon(y)dy \right| \\ &\leqslant \int_{B_\delta} |w(x-y) - w(x)|\rho_\varepsilon(y)dy + \int_{\mathbb{R}^N - B_\delta} |w(x-y) - w(x)|\rho_\varepsilon(y)dy \\ &\leqslant \int_{B_\delta} |w(x-y) - w(x)|\rho_\varepsilon(y)dy \\ &\quad + 2\sup_{z\in\mathbb{R}^N} |w(z)| \int_{\mathbb{R}^N - B_\delta} \rho_\varepsilon(y)dy \leqslant \eta \quad \text{对所有 } \varepsilon \leqslant \varepsilon_0. \end{aligned}$$

这就得到

$$\sup_{x\in\mathbb{R}^N}|(w\star\rho_\varepsilon)(x)-w(x)|\to 0 \quad \text{当 } \varepsilon\to 0 \text{ 时},$$

此因 $\eta>0$ 是任意的. (a) 证毕.

(ii) 设函数 $v\in L^1(\mathbb{R}^N)$ 给定. 则对每个 $\varepsilon>0$, 由 Fubini 定理 (定理 1.15-5)

$$\begin{aligned}\int_{\mathbb{R}^N}|(v\star\rho_\varepsilon)|(x)dx &\leqslant \int_{\mathbb{R}^N}\left(\int_{\mathbb{R}^N}|v(x-y)|\rho_\varepsilon(y)dy\right)dx\\ &=\int_{\mathbb{R}^N}\left(\int_{\mathbb{R}^N}|v(x-y)|dx\right)\rho_\varepsilon(y)dy.\end{aligned}$$

所以 $v\star\rho_\varepsilon\in L^1(\mathbb{R}^N)$ 且

$$\|v\star\rho_\varepsilon\|_{L^1(\mathbb{R}^N)}\leqslant\|v\|_{L^1(\mathbb{R}^N)} \quad \text{对每个 } \varepsilon>0.$$

由于 $\overline{\mathcal{D}(\mathbb{R}^N)}=L^1(\mathbb{R}^N)$ (定理 2.6-2), 故存在函数 $v_k\in\mathcal{D}(\mathbb{R}^N)\subset L^1(\mathbb{R}^N)$ 的序列 $(v_k)_{k=1}^\infty$ 使得当 $k\to\infty$ 时 $\|v_k-v\|_{L^1(\mathbb{R}^N)}\to 0$. 此外, 给定一开子集 $V\subset\subset\mathbb{R}^N$ 及任一整数 $k\geqslant 1$,

$$\begin{aligned}\|v\star\rho_\varepsilon-v\|_{L^1(V)}&\leqslant\|(v-v_k)\star\rho_\varepsilon\|_{L^1(V)}+\|v_k\star\rho_\varepsilon-v_k\|_{L^1(V)}+\|v-v_k\|_{L^1(V)}\\ &\leqslant 2\|v-v_k\|_{L^1(\mathbb{R}^N)}+\left(\int_V dx\right)\sup_{x\in\mathbb{R}^N}|(v_k\star\rho_\varepsilon)(x)-v_k(x)|.\end{aligned}$$

所以, 对于任意的 $k\geqslant 1$,

$$\limsup_{\varepsilon\to 0}\|v\star\rho_\varepsilon-v\|_{L^1(V)}\leqslant 2\|v-v_k\|_{L^1(\mathbb{R}^N)}.$$

此因根据 (a) $\lim_{\varepsilon\to 0}\sup_{x\in\mathbb{R}^N}|(v_k\star\rho_\varepsilon)(x)-v_k(x)|=0$ (作为空间 $\mathcal{D}(\mathbb{R}^N)$ 中的函数, 每一个函数 v_k 都是有界且一致连续的). 由于选取 k 充分大, 可使 $\|v-v_k\|_{L^1(\mathbb{R}^N)}$ 任意地小, 因此

$$\limsup_{\varepsilon\to 0}\|v\star\rho_\varepsilon-v\|_{L^1(V)}=0.$$

(b) 得证. □

我们现在已经能够来证明 Δ 对于 $L^1_{\mathrm{loc}}(\Omega)$ 中函数的次椭圆性. 要注意, 在下述证明的 (iii) 引入的函数 $E\in L^1_{\mathrm{loc}}(\mathbb{R}^N)$ 正是 Laplace 方程的基本解 (习题 6.3-4), 而在 (ii) 引入的函数 $E_\varepsilon,\varepsilon>0$, 当 $\varepsilon\to 0$ 时在 $\mathbb{R}^N$ 中几乎处处收敛于 E, 但与 E 不同的是它们在原点不再有奇性.

定理 6.4-2 (Weyl 引理[8]: Δ 的次椭圆性) *设 Ω 是 $\mathbb{R}^N$ 的一个开子集, 而 $v\in L^1_{\mathrm{loc}}(\Omega)$ 及 $f\in\mathcal{C}^m(\Omega), m\geqslant 0$ 是某一整数, 是两个满足下式的函数*

$$\int_\Omega v\Delta\varphi dx=\int_\Omega f\varphi dx \quad \text{对所有 } \varphi\in\mathcal{D}(\Omega).$$

[8] H. Weyl [1940]: The method of orthogonal projection in potential theory. Duke Mathematical Journal **7**, 414–444.

则 $v \in \mathcal{C}^m(\Omega)$. 所以如果 $f \in \mathcal{C}^\infty(\Omega)$, 则 $v \in \mathcal{C}^\infty(\Omega)$.

证明 设 ω_N 表示 $\mathbb{R}^N$ 中单位球的体积. 对每个 $\varepsilon > 0$, 如下定义函数 $E_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^N)$:

$$E_\varepsilon(x) := -\frac{1}{4\omega_2}\ln(|x|^2+\varepsilon) \quad \text{对每个 } x \in \mathbb{R}^2, \text{若 } N=2,$$
$$E_\varepsilon(x) := \frac{1}{N(2-N)\omega_N}(|x|^2+\varepsilon)^{\frac{2-N}{2}} \quad \text{对每个 } x \in \mathbb{R}^N, \text{若 } N \geqslant 3.$$

(i) 对每个 $\varepsilon > 0$, 由

$$\rho_\varepsilon := \Delta E_\varepsilon$$

定义的函数 $\rho_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}^N)$ 满足

$$\rho_\varepsilon(x) \geqslant 0 \quad \text{对所有 } x \in \mathbb{R}^N, \quad \int_\Omega \rho_\varepsilon(y)dy = 1,$$

$$\text{对每个 } \delta > 0, \quad \int_{\mathbb{R}^N - B_\delta} \rho_\varepsilon(y)dy \to 0 \quad \text{当 } \varepsilon \to 0 \text{ 时}, \quad \text{其中 } B_\delta := B(0;\delta).$$

首先, 关系式

$$\Delta E_\varepsilon(x) = \frac{\varepsilon}{\omega_N(|x|^2+\varepsilon)^{\frac{N}{2}+1}} \quad \text{对每个 } x \in \mathbb{R}^N,$$

说明对所有 $x \in \mathbb{R}^N, \rho_\varepsilon(x) \geqslant 0$. 其次, 熟知的公式

$$\int_{\mathbb{R}^N} F(|x|)dx = N\omega_N \int_0^\infty F(r)r^{N-1}dr$$

对任何可测函数 $F : [0,\infty[\to [0,\infty[$ 成立, 将其用于上述函数, 得

$$\begin{aligned}\int_{\mathbb{R}^N} \Delta E_\varepsilon(y)dy &= N\varepsilon \int_0^\infty (r^2+\varepsilon)^{-\frac{N}{2}-1} r^{N-1} dr \\ &= \int_0^\infty \frac{d}{dr}[(1+\varepsilon r^{-2})^{-\frac{N}{2}}]dr = 1.\end{aligned}$$

所以 $\int_\Omega \rho_\varepsilon(y)dy = 1$. 最后, 对每个 $\delta > 0$ 都成立公式

$$\int_{\mathbb{R}^N - B_\delta} F(|x|)dx = N\omega_N \int_\delta^\infty F(r)r^{N-1}dr,$$

类似地给出

$$\begin{aligned}\int_{\mathbb{R}^N - B_\delta} \Delta E_\varepsilon(x)dx &= N\varepsilon \int_\delta^\infty (r^2+\varepsilon)^{-\frac{N}{2}-1} r^{N-1} dr \\ &= \int_\delta^\infty \frac{d}{dr}[(1+\varepsilon r^{-2})^{-\frac{N}{2}}]dr = 1-(1+\varepsilon\delta^{-2})^{-N/2},\end{aligned}$$

这说明, 对每个 $\delta > 0$, 当 $\varepsilon \to 0$ 时, $\int_{\mathbb{R}^N - B_\delta} \rho_\varepsilon(y)dy \to 0$.

(ii) 给定一个函数 $v \in L^1_{\text{loc}}(\Omega)$ 及一个开子集 $V \subset\subset \Omega$, 如下定义函数 $\widetilde{v} \in L^1(\mathbb{R}^N)$:

$$\widetilde{v}(x) := v(x) \quad 若\ x \in V; \quad \widetilde{v}(x) := 0 \quad 若\ x \in \mathbb{R}^N - V.$$

则存在序列 $(\varepsilon(k))_{k=1}^{\infty}$ 使得对所有 $k \geqslant 1, \varepsilon(k) > 0$, 当 $k \to \infty$ 时 $\varepsilon(k) \to 0$, 而且对几乎所有 $x \in V$,

$$\begin{aligned} v(x) &= \lim_{k\to\infty} \int_{\mathbb{R}^N} \widetilde{v}(x-y)\Delta E_{\varepsilon(k)}(y)dy \\ &= \lim_{k\to\infty} \int_{\mathbb{R}^N} \widetilde{v}(y)\Delta E_{\varepsilon(k)}(x-y)dy. \end{aligned}$$

因为函数 $\rho_\varepsilon = \Delta E_\varepsilon, \varepsilon > 0$, 满足定理 6.4-1 的所有假设, 将该定理的 (b) 应用于函数 $\widetilde{v} \in L^1(\mathbb{R}^N)$, 有

$$\|\widetilde{v} \star \Delta E_\varepsilon - \widetilde{v}\|_{L^1(V)} \to 0 \quad 当\ \varepsilon \to 0\ 时.$$

因此存在函数族 $(\widetilde{v} \star \Delta E_\varepsilon)_{\varepsilon>0}$ 的一个子序列 $(\widetilde{v} \star \Delta E_{\varepsilon(k)})_{k=1}^{\infty}$ 几乎处处收敛于函数 $\widetilde{v}|_V = v|_V \in L^1(V)$.

(iii) 设 U 与 V 是 $\mathbb{R}^N$ 的两个开子集, 满足 $U \subset\subset V \subset\subset \Omega$, 又设 $\delta = \inf_{x\in U} d(x, \mathbb{R}^N - V) > 0$, 而函数 $\alpha \in \mathcal{D}(\mathbb{R}^N)$ 选得使对某个 $0 < \delta_1 < \delta$, 在 B_{δ_1} 中 $\alpha = 1$ 且 $\alpha|_{B_\delta} \in \mathcal{D}(B_\delta)$. 最后, 设函数 $E \in L^1_{\text{loc}}(\mathbb{R}^N)$ 在 $\mathbb{R}^N$ 中几乎处处定义如下

$$\begin{aligned} E(x) &:= \frac{1}{2\omega_2} \ln|x| \quad 对所有\ x \neq 0,\ 若\ N = 2, \\ E(x) &:= \frac{1}{N(2-N)\omega_N}|x|^{2-N} \quad 对所有\ x \neq 0,\ 若\ N \geqslant 3. \end{aligned}$$

设函数 $v \in L^1_{\text{loc}}(\Omega)$ 及函数 $f \in \mathcal{C}^m(\Omega)$, $m \geqslant 0$ 是某整数, 是给定的且满足

$$\int_\Omega v\Delta\varphi dx = \int_\Omega f\varphi dx \quad 对所有\ \varphi \in \mathcal{D}(\Omega).$$

则

$$v(x) = \int_{B_\delta} f(x-y)(\alpha E)(y)dy + \int_V v(y)[\Delta((1-\alpha)E)](x-y)dy \quad 对几乎所有\ x \in U.$$

为符号简单起见, 记 $E_k := E_{\varepsilon(k)}, k \geqslant 1$. 在 (ii) 中我们证明了

$$v(x) = \lim_{k\to\infty} \int_{\mathbb{R}^N} \widetilde{v}(y)\Delta E_k(x-y)dy \quad 对几乎所有\ x \in V,$$

其中 $\widetilde{v}(x) := v(x)$ 若 $x \in V$, $\widetilde{v}(x) = 0$ 若 $x \in \mathbb{R}^N - V$. 我们现在证明, 几乎对所有 $x \in U$, 当 $k \to \infty$ 时这个极限实际就由上面所宣示的表达式给出. 为此, 设函

数 $\alpha \in \mathcal{D}(\mathbb{R}^N)$ 及 $E \in L^1_{\text{loc}}(\mathbb{R}^N)$ 定义如前, 而 x 是开集 V 中的一点. 对每个整数 $k \geqslant 1$, 我们可将积分写为

$$\begin{aligned}\int_{\mathbb{R}^N} \widetilde{v}(y)\Delta E_k(x-y)dy &= \int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta(\alpha E_k + (1-\alpha)E_k)](x-y)dy \\ &= \int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta(\alpha E_k)](x-y)dy + \int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta((1-\alpha)E_k)](x-y)dy.\end{aligned}$$

然后分别讨论上面两个积分中的每一个当 $k \to \infty$ 时的性态.

首先, 我们注意对每一个 $k \geqslant 1$, 出现在第一个积分中的函数

$$\varphi_k : y \in \mathbb{R}^N \to \varphi_k(y) := (\alpha E_k)(x-y)$$

是 $\mathcal{C}^\infty$ 类的 (函数 α 与 E_k 都是 $\mathcal{C}^\infty$ 类的), 而且 $\operatorname{supp} \varphi_k \subset B_\delta(x) := \{y \in \mathbb{R}^N; |x-y| < \delta\}$, 此因 $\alpha|_{B_\delta} \in \mathcal{D}(B_\delta)$. 此外, 由 δ 的定义, $B_\delta(x) \subset V \subset \Omega$, 故 $\varphi_k|_\Omega \in \mathcal{D}(\Omega)$. 这样, 由假定即得

$$\begin{aligned}\int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta(\alpha E_k)](x-y)dy &= \int_\Omega v(y)\Delta\varphi_k(y)dy \\ &= \int_\Omega f(y)\varphi_k(y)dy = \int_\Omega f(y)[\alpha E_k](x-y)dy \\ &= \int_{B_\delta(x)} f(y)[\alpha E_k](x-y)dy = \int_{B_\delta} f(x-y)[\alpha E_k](y)dy.\end{aligned}$$

因为 $|E_k(y)| \leqslant |E(y)|$ 对所有 $y \in B_\delta - \{0\}$ 成立 (若 $N = 2$, 不失一般性我们可以假定 k 充分大而 δ 充分小, 从而保证对所有 $x \in B_\delta, 0 < |x|^2 + \varepsilon(k) \leqslant 1$) 以及 $E \in L^1(B_\delta)$, 应用 Lebesgue 控制收敛定理 (定理 1.15-3), 可以得到

$$\begin{aligned}\lim_{k\to\infty} \int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta(\alpha E_k)](x-y)dy &= \lim_{k\to\infty} \int_{B_\delta} f(x-y)[\alpha E_k](y)dy \\ &= \int_{B_\delta} f(x-y)[\alpha E](y)dy.\end{aligned}$$

其次, 我们注意出现在第二个积分中的函数 $\Delta((1-\alpha)E_k)$ 可以展开为

$$\Delta((1-\alpha)E_k) = -(\Delta\alpha)E_k - 2\nabla\alpha \cdot \nabla E_k + (1-\alpha)\Delta E_k,$$

其中 $\nabla v := (\partial_i v)_{i=1}^N$ 对任何足够光滑函数 $v : \Omega \to \mathbb{R}^N$. 考虑到

$$\operatorname{supp} \Delta\alpha \subset B_\delta - B_{\delta_1}, \quad \operatorname{supp} \nabla\alpha \subset B_\delta - B_{\delta_1} \quad \text{及} \quad \operatorname{supp}(1-\alpha) \subset \mathbb{R}^N - B_{\delta_1}.$$

我们可以推导得

$$\int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta((1-\alpha)E_k)](x-y)dy = \int_{\mathbb{R}^N} \widetilde{v}(x-y)[\Delta((1-\alpha)E_k)](x-y)dy$$

$$
\begin{aligned}
&= -\int_{B_\delta - B_{\delta_1}} \widetilde{v}(x-y)[(\Delta\alpha)E_k](y)dy - 2\int_{B_\delta - B_{\delta_1}} \widetilde{v}(x-y)[\nabla\alpha\cdot\nabla E_k](y)dy \\
&\quad + \int_{A\delta_1(x)} \widetilde{v}(x-y)\Delta E_k(y)dy,
\end{aligned}
$$

其中集合 $A_{\delta_1}(x) := (\mathbb{R}^N - B_{\delta_1})\bigcap\{y\in\mathbb{R}^N; (y-x)\in V\}$ 是有界的. 因为

$$
|E_k(y)| \leqslant |E(y)| = \frac{1}{2\pi}|\ln|y|| \leqslant \frac{1}{2\pi}|\ln\delta_1| \quad \text{对所有 } y\in B_\delta - B_{\delta_1}, \text{若 } N=2
$$

(如前面已说明过的),

$$
\begin{aligned}
|E_k(y)| &\leqslant |E(y)| \leqslant \frac{1}{N(N-2)\omega_N\delta_1^{N-2}} \quad \text{对所有 } y\in B_\delta - B_{\delta_1}, \text{若 } N\geqslant 3, \\
\|\nabla E_k(y)\| &\leqslant \frac{1}{N\omega_N}\ \frac{|y|}{(|y|^2+\varepsilon)^{\frac{N}{2}+1}} \\
&\leqslant \frac{1}{N\omega_N}\ \frac{1}{(|y|^2+\varepsilon)^{\frac{N}{2}}} \leqslant \frac{1}{N\omega_N\delta_1^N} \quad \text{对所有 } y\in B_\delta - B_{\delta_1}, \\
|\Delta E_k(y)| &= \frac{\varepsilon}{\omega_N(|y|^2+\varepsilon)^{\frac{N}{2}+1}} \leqslant \frac{1}{\omega_N(|y|^2+\varepsilon)^{\frac{N}{2}}} \leqslant \frac{1}{\omega_N\delta_1^N} \quad \text{对所有 } y\in\mathbb{R}^N - B_{\delta_1},
\end{aligned}
$$

再利用 Lebesgue 控制收敛定理, 可证得

$$
\begin{aligned}
&\lim_{k\to\infty}\int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta((1-\alpha)E_k)](y)dy \\
&= -\int_{B_\delta - B_{\delta_1}} \widetilde{v}(x-y)[(\Delta\alpha)E](y)dy \\
&\quad -2\int_{B_\delta - B_{\delta_1}} \widetilde{v}(x-y)[\nabla\alpha\cdot\nabla E](y)dy + \int_{A\delta_1(x)} \widetilde{v}(x-y)\Delta E(y)dy \\
&= \int_{\mathbb{R}^N} \widetilde{v}(x-y)[\Delta((1-\alpha)E)](y)dy = \int_{\mathbb{R}^N} \widetilde{v}(y)[\Delta((1-\alpha)E)](x-y)dy \\
&= \int_V v(y)[\Delta((1-\alpha)E)](x-y)dy.
\end{aligned}
$$

(iv) 结论. 我们所要确立的性质, 即 $v\in\mathcal{C}^m(\Omega)$ 如果 $f\in\mathcal{C}^m(\Omega)$, 是局部的, 只需证明对任意开集 $U\subset\subset\Omega$, $v\in\mathcal{C}^m(U)$ (按通常不规范的说法, 这意味着, 作为几乎处处相等的函数等价类, $v\in L^1(U)$ 包含一个 U 中的 $\mathcal{C}^m$ 类函数).

$f\in\mathcal{C}^m(\Omega)$ 的假设意味着函数

$$
x\in U\to\int_{B_\delta} f(x-y)(\alpha E)(y)dy
$$

在开集 U 上是 $\mathcal{C}^m$ 类的; 此外, 函数

$$
x\in U\to\int_V v(y)[\Delta((1-\alpha)E)](x-y)dy
$$

在 U 上是 $\mathcal{C}^\infty$ 类的 (这种可微性质将在定理 7.4-1 中用独立的推导来证明). 这样, 在 (iii) 中确定的关系式

$$v(x) = \int_{B_\delta} f(x-y)(\alpha E)(y)dy + \int_V v(y)[\Delta((1-\alpha)E)](x-y)dy \quad \text{对几乎所有 } x \in U$$

说明 $v \in \mathcal{C}^m(U)$. □

习题

6.4-1 对定理 6.4-2 当 $N=1$ 时, Ω 是 $\mathbb{R}$ 的一个开区间的情形, 给出一个证明 (书中给出的证明适用于维数 $N \geqslant 2$ 的情形).

6.5 Sobolev 空间 $W^{m,p}(\Omega)$ 及 $H^m(\Omega)$: 基本性质

诚如本章中将充分展示出的, Sobolev 空间 $H^m(\Omega)$ 与 $H_0^m(\Omega)$, 其中 $m \geqslant 1$ 是整数, 在线性椭圆边值问题以及诸如障碍问题等一些 "柔弱的非线性" 问题的分析中起着关键的作用. 作为更一般的 Sobolev 空间 $W^{m,p}(\Omega)$, 其中 p 为满足 $1 \leqslant p \leqslant \infty$ 的任意扩展的实数, $p=2$ 的特殊情况, 空间 $H^m(\Omega)$ 具有独特的性质: 是 Hilbert 空间 (见定理 6.5-1).

为了避免在以后的章节中重复, 即使如前面所述在本章中只需 $p=2$ 的特殊情况, 我们在本节及下一节中还是给出更一般的空间 $W^{m,p}(\Omega)$ 的基本性质.

尽管本节所讨论的性质在关于集合 Ω 很弱的假设下都成立, Ω 可以是 $\mathbb{R}^N$ 的任意开子集, 也可以是只有 "有限宽度" 的, 然而在下一节中开集 Ω 将被假设为有界的且具有 Lipschitz 连续的边界 (1.18 节).

它们的对偶空间将在本章稍后 (6.11 节) 予以讨论.

设 Ω 是 $\mathbb{R}^N$ 的任一开子集. 对于每个整数 $m \geqslant 1$ 及每个扩展的实数 $1 \leqslant p \leqslant \infty$, (实) Sobolev 空间[9)]

$$W^{m,p}(\Omega) \quad \text{或} \quad H^m(\Omega) \quad \text{若 } p=2,$$

---

9) 冠名源自:

S. L. Sobolev [1938]: On a theorem of functional analysis. Matematicheskii Sbornik **46**, 471–496 (原文系俄文论文).

S. L. Sobolev [1950]: Applications of Functional Analysis in Mathematical Physics. Leningrad (俄文版; 英译本: American Mathematical Society, 1963). (中译本: 王柔怀等译, 泛函分析在数学物理中的应用. 北京: 科学出版社, 1959. —— 译者注).

空间 $H^1(\Omega)$ 的定义 (用称之为 "quasi-dérivées" 的弱导数) 及其一些基本性质, 也可在下文的 205 页中找到:

J. Leray [1933]: Sur le mouvement d'un liquide visqueux emplissant l'espace. Acta Mathematica **63**, 193–248.

由如下函数 $v \in L^p(\Omega)$ 组成: 对所有重指标 $\boldsymbol{\alpha}, 1 \leqslant |\boldsymbol{\alpha}| \leqslant m, v$ 的弱偏导数 $\partial^{\boldsymbol{\alpha}} v$ 也属于 $L^p(\Omega)$. 根据弱偏导数的定义 (见 6.3 节; 注意这是由于 $L^p(\Omega) \subset L^1_{\text{loc}}(\Omega)$ 对任何 $1 \leqslant p \leqslant \infty$), 一个函数 $v \in L^p(\Omega)$ 属于 $W^{m,p}(\Omega)$, 如果对每个重指标 $\boldsymbol{\alpha}, 1 \leqslant |\boldsymbol{\alpha}| \leqslant m$, 存在一个函数 $\partial^{\boldsymbol{\alpha}} v \in L^p(\Omega)$ 使得

$$\int_\Omega (\partial^{\boldsymbol{\alpha}} v)\varphi dx = (-1)^{|\boldsymbol{\alpha}|} \int_\Omega v \partial^{\boldsymbol{\alpha}} \varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega).$$

我们已经知道, 这样一个函数 $\partial^{\boldsymbol{\alpha}} v \in L^p(\Omega)$ 是由上述关系式唯一确定的, 而且如果还有 $v \in \mathcal{C}^m(\Omega)$, 那么 $\partial^{\boldsymbol{\alpha}} v$ 就是通常的偏导数 (定理 6.3-3).

要注意, 每一个空间 $W^{m,p}(\Omega), m \geqslant 1$, 都被定义为 $L^p(\Omega)$ 的一个子空间, 实际上是严格地包含在 $L^p(\Omega)$ 内的 (习题 6.5-1).

我们现在开始给出下面需要的 Sobolev 空间的各种基本性质, 但当我们在本节及下一节着手进行时, 会发现有越来越多的假设加在开集 Ω 上 (然而为简单起见, 我们没必要陈述使每一个定理成立的 "最弱的" 合理假设). 回忆一下, 一个赋范向量空间是自反的, 如果它能利用特定的等距映射等同于其对偶空间的对偶空间 (5.14 节).

定理 6.5-1 设 Ω 是 $\mathbb{R}^N$ 的开子集, 而 $m \geqslant 1$ 是一整数. 装备以范数

$$\begin{aligned}
v \to \|v\|_{m,p,\Omega} &:= \left(\int_\Omega \sum_{|\boldsymbol{\alpha}| \leqslant m} |\partial^{\boldsymbol{\alpha}} v|^p dx \right)^{1/p} \\
&= \left(\sum_{0 \leqslant |\boldsymbol{\alpha}| \leqslant m} \|\partial^{\boldsymbol{\alpha}} v\|^p_{L^p(\Omega)} \right)^{1/p} \qquad \text{若 } 1 \leqslant p < \infty, \\
v \to \|v\|_{m,\infty,\Omega} &:= \max_{|\boldsymbol{\alpha}| \leqslant m} \|\partial^{\boldsymbol{\alpha}} v\|_{L^\infty(\Omega)} \quad \text{若 } p = \infty, \\
v \to \|v\|_{m,\Omega} &:= \left(\int_\Omega \sum_{|\boldsymbol{\alpha}| \leqslant m} |\partial^{\boldsymbol{\alpha}} v|^2 dx \right)^{1/2} \\
&= \left(\sum_{0 \leqslant |\boldsymbol{\alpha}| \leqslant m} \|\partial^{\boldsymbol{\alpha}} v\|^2_{L^2(\Omega)} \right)^{1/2} \qquad \text{若 } p = 2,
\end{aligned}$$

的 Sobolev 空间 $W^{m,p}(\Omega)$ 是 Banach 空间.

空间 $W^{m,p}(\Omega)$ 当 $1 \leqslant p < \infty$ 时是可分的, 当 $1 < p < \infty$ 时是自反的.

空间 $H^m(\Omega) = W^{m,2}(\Omega)$ 是 Hilbert 空间.

证明 容易验证, 对每个 $1 \leqslant p \leqslant \infty$, 映射 $\|\cdot\|_{m,p,\Omega}$ 是 $W^{m,p}(\Omega)$ 上的范数.

设 $1 \leqslant p \leqslant \infty, (v_k)_{k=1}^\infty$ 是装备以这种范数的 $W^{m,p}(\Omega)$ 中的 Cauchy 序列. 因为对每个 $0 \leqslant |\boldsymbol{\alpha}| \leqslant m$, 有

$$\|\partial^{\boldsymbol{\alpha}} v_k - \partial^{\boldsymbol{\alpha}} v_\ell\|_{L^p(\Omega)} \leqslant \|v_k - v_\ell\|_{m,p,\Omega} \quad \text{对所有 } k, \ell \geqslant 1,$$

又因为 $L^p(\Omega)$ 是完备的 (定理 3.4-2), 故对每个 $1 \leqslant |\boldsymbol{\alpha}| \leqslant m$ 存在一个函数 $v^{\boldsymbol{\alpha}} \in L^p(\Omega)$ 使得当 $k \to \infty$ 时, $\|\partial^{\boldsymbol{\alpha}} v_k - v^{\boldsymbol{\alpha}}\|_{L^p(\Omega)} \to 0$, 而且存在函数 $v \in L^p(\Omega)$ 使得当 $k \to \infty$ 时, $\|v_k - v\|_{L^p(\Omega)} \to 0$.

设 $\varphi \in \mathcal{D}(\Omega)$ 是一给定函数, 这就有

$$\int_\Omega (\partial^{\boldsymbol{\alpha}} v_k)\varphi dx = (-1)^{|\boldsymbol{\alpha}|}\int_\Omega v_k \partial^{\boldsymbol{\alpha}}\varphi dx \quad \text{对所有 } 1 \leqslant |\boldsymbol{\alpha}| \leqslant m \text{ 及所有 } k \geqslant 1,$$

此因 $v_k \in W^{m,p}(\Omega)$. 当 $k \to \infty$ 时, 在下述不等式中取极限

$$\left|\int_\Omega (\partial^{\boldsymbol{\alpha}} v_k)\varphi dx - \int_\Omega v^{\boldsymbol{\alpha}}\varphi dx\right| \leqslant \|\partial^{\boldsymbol{\alpha}} v_k - v^{\boldsymbol{\alpha}}\|_{L^p(\Omega)}\|\varphi\|_{L^q(\Omega)} \quad \text{对所有 } 1 \leqslant |\boldsymbol{\alpha}| \leqslant m,$$

$$\left|\int_\Omega v_k \partial^{\boldsymbol{\alpha}}\varphi dx - \int_\Omega v\partial^{\boldsymbol{\alpha}}\varphi dx\right| \leqslant \|v_k - v\|_{L^p(\Omega)}\|\partial^{\boldsymbol{\alpha}}\varphi\|_{L^q(\Omega)},$$

其中 q 表示 p 的共轭指数, 就证明了

$$\int_\Omega (v^{\boldsymbol{\alpha}})\varphi dx = (-1)^{|\boldsymbol{\alpha}|}\int_\Omega v\partial^{\boldsymbol{\alpha}}\varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega).$$

这就说明, 对每个 $|\boldsymbol{\alpha}| \leqslant m, v^{\boldsymbol{\alpha}} \in L^p(\Omega)$ 是 $v \in L^p(\Omega)$ 的 $\boldsymbol{\alpha}$ 阶弱偏导数; 因此 $v \in W^{m,p}(\Omega)$. 此外, 函数 $v^{\boldsymbol{\alpha}} \in L^p(\Omega)$ 以及范数 $\|\cdot\|_{m,p,\Omega}$ 的定义说明当 $k \to \infty$ 时 $\|v_k - v\|_{m,p,\Omega} \to 0$. 故空间 $(W^{m,p}(\Omega), \|\cdot\|_{m,p,\Omega})$ 是完备的.

为了验证当 $1 \leqslant p < \infty$ 时空间 $W^{m,p}(\Omega)$ 是可分的, 我们只讨论 $m = 1$ 的情况即可, 因 $m \geqslant 2$ 的情况的处理是类似的. 由于空间 $W^{1,p}(\Omega)$ 作为赋范向量空间可以等同于装备以积范数的积空间 $(L^p(\Omega))^{N+1}$ 的子空间

$$\left\{(v_0, v_1, \cdots, v_N) \in (L^p(\Omega))^{N+1}; \int_\Omega v_i\varphi dx = -\int_\Omega v_0\partial_i\varphi dx \text{ 对所有 } \varphi \in \mathcal{D}(\Omega), 1 \leqslant i \leqslant N\right\},$$

故 $W^{m,p}(\Omega)$ 的可分性由 $(L^p(\Omega))^{N+1}$ 的可分性 (这可由 $L^p(\Omega)$ 的可分性得到; 见定理 2.5-4) 以及可分距离空间的任一子集也是可分的 (定理 1.10-3) 这一事实得到.

由于上述 $(L^p(\Omega))^{N+1}$ 的子空间显然是闭的, 当 $1 < p < \infty$ 时 $W^{m,p}(\Omega)$ 的自反性可由定理 5.14-2(c) (它断言, 自反空间的任一闭子空间也是自反的) 以及 $L^p(\Omega), 1 < p < \infty$, 的自反性 (定理 5.14-2(e)) 得到, 因后者显然意味着 $(L^p(\Omega))^{N+1}$ 的自反性.

很容易验证, 由下式定义的双线性映射 $(\cdot,\cdot)_{m,\Omega} : H^m(\Omega) \times H^m(\Omega) \to \mathbb{R}$

$$(u, v)_{m,\Omega} := \sum_{|\boldsymbol{\alpha}| \leqslant m}\int_\Omega \partial^{\boldsymbol{\alpha}} u\partial^{\boldsymbol{\alpha}} v dx \quad \text{对所有 } u, v \in H^m(\Omega)$$

具有在空间 $H^m(\Omega)$ 上内积的所有性质, 而且 $\|v\|_{m,\Omega} = \sqrt{(v,v)_{m,\Omega}}$ 对所有 $v \in H^m(\Omega)$ 成立, 所以 $(H^m(\Omega), \|\cdot\|_{m,\Omega})$ 是 Hilbert 空间. □

注 在习题 6.11-2 中, 对空间 $W^{m,p}(\Omega), 1 < p < \infty$, 的自反性提出了一个不同的证明.

为方便起见, 在空间 $W^{m,p}(\Omega)$ 以及范数 $\|\cdot\|_{m,p,\Omega}$ 的上述定义中, 允许 $m=0$, 令

$$W^{0,p}(\Omega):=L^p(\Omega) \quad \text{且} \quad \|\cdot\|_{0,p,\Omega}:=\|\cdot\|_{L^p(\Omega)} \quad \text{若 } 1\leqslant p\leqslant\infty,$$
$$H^0(\Omega):=L^2(\Omega) \quad \text{且} \quad \|\cdot\|_{0,\Omega}:=\|\cdot\|_{L^2(\Omega)} \quad \text{若 } p=2.$$

尽管当 $1\leqslant p<\infty$ 时, 空间 $\mathcal{D}(\Omega)$ 在空间 $L^p(\Omega)$ 中稠密 (定理 2.6-2), 但空间 $\mathcal{D}(\Omega)$ 在空间 $W^{m,p}(\Omega)$ 中, $m\geqslant 1$, 并不稠密, 除非集合 $\mathbb{R}^N-\Omega$ "非常小". 例如, 可以证明, 如果 $1<p<\infty$, 则 $\overline{\mathcal{D}(\Omega)}=W^{m,p}(\Omega)$ 的必要 (但不是充分) 条件是 $\mathbb{R}^N-\Omega$ 的 Lebesgue 测度为零[10] (关于这方面的问题, 也可参阅习题 6.5-2 及 6.5-3). 上述探讨激发出下述定义.

设 Ω 是 $\mathbb{R}^N$ 中的开子集. 对每一整数 $m\geqslant 1$ 与每个实数 $1\leqslant p<\infty$, Sobolev 空间

$$W_0^{m,p}(\Omega) \quad \text{或} \ H_0^m(\Omega) \quad \text{若 } p=2,$$

定义为空间 $\mathcal{D}(\Omega)$ 在空间 $(W^{m,p}(\Omega),\|\cdot\|_{m,p,\Omega})$ 中的闭包. 由这个定义及定理 6.5-1 和定理 5.14-2(c) 立即可得, 对每个整数 $m\geqslant 1$, 若 $1\leqslant p<\infty$, 空间 $W_0^{m,p}(\Omega)$ 是可分 Banach 空间, 若 $1<p<\infty$, 它还是自反的, 而空间 $H_0^m(\Omega)$ 是 Hilbert 空间.

空间 $W_0^{m,p}(\Omega)$ 的另一些基本性质将在下一定理中予以证明 (作为补充也可参见习题 6.5-3 到 6.5-5), 其中要用到下述半范数, 对每个整数 $m\geqslant 1$ 及每个实数 $1\leqslant p<\infty$,

$$v\to|v|_{m,p,\Omega}:=\left(\int_\Omega\sum_{|\boldsymbol{\alpha}|=m}|\partial^{\boldsymbol{\alpha}}v|^p dx\right)^{1/p}=\left(\sum_{|\boldsymbol{\alpha}|=m}\|\partial^{\boldsymbol{\alpha}}v\|_{L^p(\Omega)}^p\right)^{1/p},$$
$$v\to|v|_{m,\Omega}:=\left(\int_\Omega\sum_{|\boldsymbol{\alpha}|=m}|\partial^{\boldsymbol{\alpha}}v|^2 dx\right)^{1/2}=\left(\sum_{|\boldsymbol{\alpha}|=m}\|\partial^{\boldsymbol{\alpha}}v\|_{L^2(\Omega)}^2\right)^{1/2} \quad \text{若 } p=2.$$

$\mathbb{R}^N$ 的一个子集如果位于 $\mathbb{R}^N$ 中的两个平行的超平面之间, 则称其具有有限宽度.

定理 6.5-2 设 Ω 为 $\mathbb{R}^N$ 的具有有限宽度的开子集.

(a) 对每个 $1\leqslant p<\infty$, 下述 Poincaré-Friedrichs 不等式[11] 成立: 存在一个常数 $c=c(\Omega,p)$ 使得

$$\|v\|_{0,p,\Omega}\leqslant c|v|_{1,p,\Omega} \quad \text{对所有 } v\in W_0^{1,p}(\Omega).$$

(b) 对每个 $m\geqslant 1$ 及 $1\leqslant p<\infty$, 半范数 $|\cdot|_{m,p,\Omega}$ 是空间 $W_0^{m,p}(\Omega)$ 上等价于范数 $\|\cdot\|_{m,p,\Omega}$ 的范数, 即存在常数 $C=C(\Omega,m,p)$ 使得

$$|v|_{m,p,\Omega}\leqslant\|v\|_{m,p,\Omega}\leqslant C|v|_{m,p,\Omega} \quad \text{对所有 } v\in W_0^{m,p}(\Omega).$$

---

[10] J. L. LIONS [1965]: Problèmes aux Limites dans les Equations aux Dérivées Partielles. Presses de l'Université de Montréal, Montréat, Que.

[11] 冠名源自 Henri Poincaré (1854—1912) 及 Kurt Otto Friedrichs (1901—1982). 前者对光滑函数建立了有关不等式, 后者将其推广到空间 $W_0^{1,p}(\Omega)$.

证明 只要对 $W_0^{1,p}(\Omega)$ 的稠密子空间 $\mathcal{D}(\Omega)$ 中的函数确立 Poincaré-Friedrichs 不等式就行了, 这是由于无论范数 $\|\cdot\|_{0,p,\Omega}$ 还是半范数 $|\cdot|_{1,p,\Omega}$ 都是空间 $W^{1,p}(\Omega)$ 上的连续函数, 这立即可由下述不等式得出

$$|\|v\|_{0,p,\Omega} - \|w\|_{0,p,\Omega}| \leqslant \|v - w\|_{0,p,\Omega} \leqslant \|v - w\|_{1,p,\Omega},$$
$$||v|_{1,p,\Omega} - |w|_{1,p,\Omega}| \leqslant |v - w|_{1,p,\Omega} \leqslant \|v - w\|_{1,p,\Omega}.$$

假设 Ω 位于垂直于向量 $(1,0,\cdots,0)$ 的两个平行的超平面之间, 而另设 $a > 0$ 使得 $\Omega \subset [-a,a] \times \mathbb{R}^{N-1}$. 给定任一函数 $v \in \mathcal{D}(\Omega)$, 在 $]-a,a[\ \times\mathbb{R}^{N-1}$ 中作零延拓后仍视为 v, 我们就有

$$v(x) = \int_{-a}^{x_1} \partial_1 v(t, x_2, \cdots, x_N) dt \quad 对所有\ x = (x_1, x_2, \cdots, x_N) \in\ ]-a,a[\ \times\mathbb{R}^{N-1}.$$

因而

$$\begin{aligned}|v(x)|^p &\leqslant \left(\int_{-a}^{x_1} |\partial_1 v(t, x_2, \cdots, x_N)| dt\right)^p \\ &\leqslant (a + x_1)^{p-1} \int_{-a}^{x_1} |\partial_1 v(x_1, x_2, \cdots, x_N)|^p dx_1 \\ &\leqslant (a + x_1)^{p-1} \int_{-a}^{a} |\partial_1 v(x_1, x_2, \cdots, x_N)|^p dx_1.\end{aligned}$$

于是

$$\int_{-a}^{a} |v(x)|^p dx_1 \leqslant \frac{(2a)^p}{p} \int_{-a}^{a} |\partial_1 v(x_1, x_2, \cdots, x_N)|^p dx_1,$$

故由 Fubini 定理 (定理 1.15-5) 得

$$\begin{aligned}\|v\|_{0,p,\Omega}^p &= \int_{\mathbb{R}^{N-1}} \left(\int_{-a}^{a} |v(x)|^p dx_1\right) dx_2 \cdots dx_N \\ &\leqslant \frac{(2a)^p}{p} \int_{\mathbb{R}^{N-1}} \left(\int_{-a}^{a} |\partial_1 v(x_1, x_2, \cdots, x_N)|^p dx_1\right) dx_2 \cdots dx_N \\ &= \frac{(2a)^p}{p} \|\partial_1 v\|_{0,p,\Omega}^p \leqslant \frac{(2a)^p}{p} |v|_{1,p,\Omega}.\end{aligned}$$

这就证明了 (a), 其中的常数 $c = \dfrac{(2a)}{p^{1/p}}$.

Poincaré-Friedrichs 不等式立即给出

$$|v|_{1,p,\Omega} \leqslant \|v\|_{1,p,\Omega} \leqslant (1 + c^p)^{1/p} |v|_{1,p,\Omega} \quad 对所有\ v \in W_0^{1,p}(\Omega),$$

这就对 $m = 1$ 的情况证明了 (b). 上述不等式进一步意味着

$$\begin{aligned}\|v\|_{2,p,\Omega}^p &= \|v\|_{1,p,\Omega}^p + |v|_{2,p,\Omega}^p \\ &\leqslant (1 + c^p) |v|_{1,p,\Omega}^p + |v|_{2,p,\Omega}^p \quad 对所有\ v \in \mathcal{D}(\Omega),\end{aligned}$$

而另外再用 Poincaré-Friedrichs 不等式, 有

$$|v|_{1,p,\Omega}^p = \sum_{i=1}^N \|\partial_i v\|_{0,p,\Omega}^p$$
$$\leqslant c^p \sum_{i=1}^N |\partial_i v|_{1,p,\Omega}^p = c^p |v|_{2,p,\Omega}^p \quad 对所有\ v \in \mathcal{D}(\Omega).$$

后两个不等式一起就证明了 (b) 在 $m=2$ 的情形. 同样的推导可对任何整数 $m \geqslant 3$ 的情形证明 (b).

在一般的情况, 设 $x \in \Omega \to \widehat{x} = a + Qx \in \mathbb{R}^N$, 其中 $a \in \mathbb{R}^N$ 而 Q 是 N 阶正交矩阵, 是一个笛卡儿坐标变换使得在这个变换下 Ω 的像 $\widehat{\Omega}$ 位于垂直于向量 $(1,0,\cdots,0)$ 的两个平行的超平面之间. 这样, 因为矩阵 Q 是正交的, 则 $\sum_{i=1}^N |\partial_i v(x)|^2 = \sum_{i=1}^N |\widehat{\partial}_i \widehat{v}(\widehat{x})|^2$ (这里符号的意义自明). 因为 Euclid 范数 $\|\cdot\|_2$ 与范数 $\|\cdot\|_p$ 在 $\mathbb{R}^N$ 中是等价的, 由 $\widehat{\Omega}$ 上的 Poincaré 不等式即得在 Ω 上的 Poincaré 不等式. □

习题

6.5-1 设 Ω 为 $\mathbb{R}^N$ 的开子集. 证明, 对任何 $1 \leqslant p \leqslant \infty, W^{1,p}(\Omega) \subsetneq L^p(\Omega)$.

提示: 如果 $N = 1$ 及 $\Omega =]0,1[$, 证明如下定义的函数 $v \in L^p(0,1)$: $v(x) = 0$ 若 $0 < x < \frac{1}{2}$, 而 $v(x) = 1$ 若 $\frac{1}{2} < x < 1$, 在 $L^p(0,1)$ 中没有一阶弱导数. 将此例改为对任意 $N \geqslant 2$ 的情形.

6.5-2 证明, 对任意整数 $m \geqslant 1$ 及任意的 $1 \leqslant p < \infty$, 空间 $\mathcal{D}(\mathbb{R}^N)$ 在空间 $W^{m,p}(\mathbb{R}^N)$ 中稠密.[12)]

6.5-3 证明, 如果 $\mathbb{R}^N$ 的开子集 Ω 是有界的, 则 $H_0^1(\Omega) \subsetneq H^1(\Omega)$.

提示: 验明 $H_0^1(\Omega)$ 在 $H^1(\Omega)$ 中的正交补里有非零函数.

6.5-4 设 $1 \leqslant p < \infty$.

(1) 证明在空间 $W^{1,p}(\mathbb{R}^N)$ 上的半范数 $|\cdot|_{1,p,\mathbb{R}^N}$ 是范数.

(2) 范数 $|\cdot|_{1,p,\mathbb{R}^N}$ 是否等价于 $W^{1,p}(\mathbb{R}^N)$ 上的范数 $\|\cdot\|_{1,p,\mathbb{R}^N}$?

6.5-5 设 Ω 是 $\mathbb{R}^N$ 的开子集. 对空间 $H_0^1(\Omega)$ 是无穷维的这一结论, 给出一个一行字的证明.

6.5-6 设 Ω 是 $\mathbb{R}^N$ 的开子集, $1 < p < \infty, q$ 表示 p 的共轭指数. 证明, 一个函数 $v \in L^p(\Omega)$ 属于空间 $W^{1,p}(\Omega)$ 的充分必要条件是存在常数 C 使得

$$\left|\int_\Omega v\partial_i\varphi dx\right| \leqslant C\|\varphi\|_{0,q,\Omega} \quad 对所有\ \varphi \in \mathcal{D}(\Omega).$$

提示: 利用 $L^q(\Omega)$ 中的 F. Riesz 表示定理 (定理 3.4-3).

---

12) 其证明可参阅, 如 ADAMS [1975, 推论 3.19] 或 ATTOUTH, BUTTAZZO & MICHAILLE [2006, 定理 5.1-3].

6.6 关于区域 Ω 的 Sobolev 空间 $W^{m,p}(\Omega)$ 和 $H^m(\Omega)$: 嵌入定理, 迹, Green 公式

在 6.5 节中描述的 Sobolev 空间 $W^{m,p}(\Omega)$ 的性质, 对 $\mathbb{R}^N$ 的任意开子集都成立 (除了 Poincaré-Friedrichs 不等式, 它建立在 Ω 具有有限宽度这一假设上). 与此形成鲜明对照的是, 本节中所描述的性质只是在关于 Ω 的某些特定的假设下, 如其有界性, 特别是其边界的光滑性等, 才成立. 这些性质的证明, 通常是冗长且多是技术性的, 不在这里给出. 有兴趣的读者可参阅在文献注释中所建议的参考文献.

容易证明, 对任何整数 $m \geqslant 1$, 空间 $W^{m,p}(\Omega)$ 严格地包含在空间 $L^p(\Omega)$ 内 (习题 6.5-1). 这实际上反映出, 任何一个 $L^p(\Omega)$ 中的函数要有在 $L^p(\Omega)$ 中的弱导数, 需要某种额外的"光滑性". 例如, 设 Ω 是 $\mathbb{R}^2$ 中的区域, 那么一个空间 $W^{1,p}(\Omega)$ 中的函数当 $p > 2$ 时必须在 $\overline{\Omega}$ 上连续; 当 $p = 2$ 时必属于任一空间 $L^q(\Omega), 1 \leqslant q < \infty$; 当 $1 \leqslant p < 2$ 时必属于空间 $L^{2p/(2-p)}(\Omega)$ (这些性质是下面定理 6.6-1 中所给出的性质的特殊情况). 然而要注意, 如果一个函数属于空间 $H^1(\Omega)$ (即若 $p = 2$), $\Omega \subset \mathbb{R}^2$, 则它不一定是连续的, 在习题 6.6-1 中给出了一个令人惊诧的示例: 一个 $H^1(\Omega)$ 中的函数在 Ω 中甚至处处不连续.

前面所谈及的包含关系 $W^{1,p}(\Omega) \subset \mathcal{C}^0(\overline{\Omega})$ 或 $W^{1,p}(\Omega) \subset L^q(\Omega)$, 其中 Ω 为 $\mathbb{R}^2$ 中的区域, 是下述定理中所陈述的嵌入的实例. 在以下定理中, 符号

$$X \hookrightarrow Y$$

意指赋范向量空间 X 连续地嵌入赋范向量空间 Y 中, 也就是说 $X \subset Y$ 而且存在一个常数 c 使得 $\|v\|_Y \leqslant c\|v\|_X$ 对所有 $v \in X$ 成立; 换言之, 恒等映射 $\iota : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ 是连续的.

回忆一下, 空间 $\mathcal{C}^{m,\lambda}(\overline{\Omega})$ 已在 1.18 节中定义, 而 $\mathbb{R}^N$ 中的区域在那里也已定义.

当解释这些连续嵌入 (或定理 6.6-3 中的紧嵌入) 时, 还是应该小心谨慎, 这是因为 Sobolev 空间的一个元素, 实际上是在 Ω 中几乎处处相等的函数的一个等价类. 例如, 嵌入 $W^{m,p}(\Omega) \hookrightarrow \mathcal{C}^{0,\lambda}(\overline{\Omega})$ 意味着存在一个常数 c 使得, 在空间 $W^{m,p}(\Omega)$ 的每一个等价类中, 有一个 (唯一的) 代表 v 属于空间 $\mathcal{C}^{0,\lambda}(\overline{\Omega})$ 且满足 $\|v\|_{\mathcal{C}^{0,\lambda}(\overline{\Omega})} \leqslant c\|v\|_{m,p,\Omega}$, 等等.

定理 6.6-1 (Sobolev 嵌入定理) 设 Ω 是 $\mathbb{R}^N$ 中的区域, $m \geqslant 1$ 为整数而 $1 \leqslant p < \infty$. 则下述连续嵌入成立:

$$W^{m,p}(\Omega) \hookrightarrow L^{p*}(\Omega), \quad \text{其中 } \frac{1}{p*} = \frac{1}{p} - \frac{m}{N}, \quad \text{若 } m < \frac{N}{p},$$

$$W^{m,p}(\Omega) \hookrightarrow L^q(\Omega) \quad \text{对所有满足 } 1 \leqslant q < \infty \text{ 的 } q, \text{ 若 } m = \frac{N}{p},$$

$$W^{m,p}(\Omega) \hookrightarrow \mathcal{C}^{0,m-N/p}(\overline{\Omega}), \quad \text{若 } \frac{N}{p} < m < \frac{N}{p} + 1,$$

$$W^{m,p}(\Omega) \hookrightarrow \mathcal{C}^{0,\lambda}(\overline{\Omega}) \quad \text{对所有满足 } 0<\lambda<1 \text{ 的 } \lambda, \text{ 若 } m=\frac{N}{p}+1,$$

$$W^{m,p}(\Omega) \hookrightarrow \mathcal{C}^{0,1}(\overline{\Omega}) \quad \text{若 } \frac{N}{p}+1<m.$$

Sobolev 嵌入定理的一个重要推论是, 保证嵌入 $W^{m,p}(\Omega) \hookrightarrow \mathcal{C}^0(\overline{\Omega})$ 成立的不等式, 即不等式 $mp > N$, 也保证 Sobolev 空间 $W^{m,p}(\Omega)$ 是一个 Banach 代数, 即一个 Banach 空间, 根据 2.15 节给出的定义, 也是代数. 在这种特殊情况下, 意味着如果 $mp > N$, 两个 $W^{m,p}(\Omega)$ 中函数的乘积也属于 $W^{m,p}(\Omega)$, 而且以这种方式定义的双线性映射 $(u,v) \in W^{m,p}(\Omega) \times W^{m,p}(\Omega) \to uv \in W^{m,p}(\Omega)$ 关于范数 $\|\cdot\|_{m,p,\Omega}$ 是连续的 (2.11 节). 更明确地, 我们有:

定理 6.6-2 ($W^{m,p}(\Omega)$ 是 Banach 代数, 若 $mp > N$) 设 Ω 是 $\mathbb{R}^N$ 中的区域, $m \geqslant 1$ 是整数, 又设 $1 \leqslant p < \infty$ 使得 $mp > N$. 则

$$u, v \in W^{m,p}(\Omega) \Rightarrow uv \in W^{m,p}(\Omega),$$

且存在常数 c 使得

$$\|uv\|_{m,p,\Omega} \leqslant c\|u\|_{m,p,\Omega}\|v\|_{m,p,\Omega} \quad \text{对所有 } u, v \in W^{m,p}(\Omega).$$

赋范向量空间 X 称为紧嵌入到赋范向量空间 Y, 如果 $X \hookrightarrow Y$ 且恒等映射 $\iota : x \in X \to \iota(x) = x \in Y$ 是紧线性算子; 或等价地, ι 将每个有界序列 $(x^k)_{k=1}^{\infty}$ 映射为序列 $(\iota(x^k))_{k=1}^{\infty}$, 它包含一个在 Y 中收敛的子列 (定理 2.10-1); 这种紧嵌入表示为

$$X \Subset Y.$$

下述结果就是定理 6.6-1 中的连续嵌入, 但它们除此之外还是紧的; 其中数 p^* 仍如定理 6.6-1 中所定义.

定理 6.6-3 (Rellich-Kondrachov 紧嵌入定理[13]) 设 Ω 为 $\mathbb{R}^N$ 中的区域, $m \geqslant 1$ 是整数, 而 $1 \leqslant p < \infty$. 则下述紧嵌入成立:

$$W^{m,p}(\Omega) \Subset L^q(\Omega) \quad \text{对所有满足 } 1 \leqslant q < p^* \text{ 的 } q, \text{ 若 } m < \frac{N}{p},$$

$$W^{m,p}(\Omega) \Subset L^q(\Omega) \quad \text{对所有满足 } 1 \leqslant q < \infty \text{ 的 } q, \text{ 若 } m = \frac{N}{p},$$

$$W^{m,p}(\Omega) \Subset C(\overline{\Omega}), \quad \text{若 } \frac{N}{p} < m.$$

[13] F. RELLICH [1930]: Ein Satz über mittlere Konvergenz. Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, 30–35.

V. I. KONDRACHOV [1945]: Certain properties of functions in the spaces L^p. Doklady Akademii Nauk SSSR **48**, 535–538 (俄文).

注意, Rellich-Kondrachov 定理意味着紧嵌入

$$W^{1,p}(\Omega) \Subset L^p(\Omega), \quad p \geqslant 1,$$

总是成立的, 即与维数 N 无关.

当 Ω 是区域时, Sobolev 空间 $W^{m,p}(\Omega)$ 中函数的另一个重要性质是, 它们可以被光滑函数逼近. 空间 $\mathcal{C}^\infty(\overline{\Omega})$ 已在 1.18 节中定义.

定理 6.6-4 (用光滑函数逼近) 设 Ω 为 $\mathbb{R}^N$ 中的区域, $m \geqslant 0$ 是整数, 而 $1 \leqslant p < \infty$. 则空间 $\mathcal{C}^\infty(\overline{\Omega})$ 在空间 $W^{m,p}(\Omega)$ 中稠密.

对于下面的结果, 为简单计, 多半只给出其在 $m = 1$ 的特殊情况, 即 Sobolev 空间 $W^{1,p}(\Omega)$ 中函数的情况, 虽然类似的性质在 Sobolev 空间 $W^{m,p}(\Omega), m \geqslant 2$, 也成立. 首先, 我们需定义一些函数空间.

设 Ω 是 $\mathbb{R}^N$ 中一区域, Γ 表示其边界, $1 \leqslant p < \infty$. 空间 $\mathcal{L}^p(\Gamma)$ 定义为满足 $|f|^p \in \mathcal{L}^1(\Gamma)$ (空间 $\mathcal{L}^1(\Gamma)$ 已在 1.18 节中定义) 的所有函数 $f : \Gamma \to \mathbb{R}$ 组成的空间. 空间 $\mathcal{L}^p(\Gamma)$ 中的函数, 模 $d\Gamma$ 几乎处处相等的等价类形成的空间表示为

$$L^p(\Gamma).$$

利用函数 $g \in \mathcal{L}^1(\Gamma)$ (1.18 节) 积分 $\int_\Gamma g d\Gamma$ 的定义, 并结合类似于定理 2.5-2 和 3.4-2 证明中所用的推导, 可以确立, 函数

$$f \in L^p(\Gamma) \to \|f\|_{L^p(\Gamma)} := \left(\int_\Gamma |f|^p d\Gamma \right)^{1/p}$$

是空间 $L^p(\Gamma)$ 上的范数, 而空间 $(L^p(\Gamma), \|\cdot\|_{L^p(\Gamma)})$ 是 Banach 空间.

设 $v : \overline{\Omega} \to \mathbb{R}$ 是一连续函数, 其中 Ω 为 $\mathbb{R}^N$ 的一个开子集. 则它在 Ω 的边界 Γ 上的迹是连续函数 $\mathrm{tr}\, v : \Gamma \to \mathbb{R}$, 其定义由 $(\mathrm{tr}\, v)(x) = v(x)$ 对所有 $x \in \Gamma$ 给出. 当 Ω 是 $\mathbb{R}^N$ 中的区域时, Sobolev 空间 $W^{1,p}(\Omega)$ 中函数的一个值得关注的性质是, 即使函数在 $\overline{\Omega}$ 上不连续, 也可以在 Γ 上定义其广义的 "迹". 这个推广的基础来自下述观察: 设 Ω 为 $\mathbb{R}^N$ 中的一个区域, 而 $1 \leqslant p < N$. 则可以证明, 映射

$$\mathrm{tr} : \mathcal{C}^\infty(\overline{\Omega}) \to L^{p^\#}(\Gamma)$$

是适定且连续的, 如果空间 $\mathcal{C}^\infty(\overline{\Omega})$ 被赋予范数 $\|\cdot\|_{1,p,\Omega}$ 而 $p^\#$ 如下面定理 6.6-5 中所定义的. 因为空间 $\mathcal{C}^\infty(\overline{\Omega})$ 在空间 $W^{1,p}(\Omega)$ 中稠密 (定理 6.6-4) 并且空间 $L^{p^\#}(\Gamma)$ 是完备的, 故存在唯一的从空间 $W^{1,p}(\Omega)$ 到空间 $L^{p^\#}(\Gamma)$ 的连续线性延拓 (定理 3.1-1), 其在子空间 $\mathcal{C}^\infty(\overline{\Omega})$ 上与经典的迹算子 tr 完全一致. 这个延拓称为迹算子, 仍将以同样的符号 tr 表示. 以这种方式定义的函数 $\mathrm{tr}\, v \in L^{p^\#}(\Gamma)$ 称为函数 $v \in W^{1,p}(\Omega)$ 的迹.

我们现在叙述迹算子的各种重要性质, 如连续性、紧性, 以及当 Ω 为区域时, 如何用迹算子为空间 $W_0^{1,p}(\Omega)$ 和 $W_0^{2,p}(\Omega)$ 提供另一个等价的定义.

定理 6.6-5 (迹算子的性质)　设 Ω 为 $\mathbb{R}^N$ 中的区域.

(a) 设 $1 \leqslant p < \infty$, 则

$$\operatorname{tr} \in \mathcal{L}(W^{1,p}(\Omega); L^{p^\#}(\Gamma)), \quad \text{其中} \frac{1}{p^\#} := \frac{1}{p} - \frac{p-1}{p(N-1)}, \quad \text{若 } 1 \leqslant p < N,$$
$$\operatorname{tr} \in \mathcal{L}(W^{1,p}(\Omega); L^q(\Gamma)), \quad \text{对所有满足 } 1 \leqslant q < \infty \text{ 的 } q, \quad \text{若 } p = N,$$
$$\operatorname{tr} \in \mathcal{L}(W^{1,p}(\Omega); \mathcal{C}(\Gamma)), \quad \text{若 } N < p.$$

(b) 若 $1 < p < N$, 则迹算子 $\operatorname{tr}: W^{1,p}(\Omega) \to L^q(\Omega)$ 对所有满足 $1 \leqslant q < p^\#$ 的 q 是紧的.

(c) 设 $1 \leqslant p < \infty$, 则定义为 $\mathcal{D}(\Omega)$ 在 $W^{1,p}(\Omega)$ 中闭包的空间 $W_0^{1,p}(\Omega)$ (6.5 节) 也可由下式给出:

$$W_0^{1,p}(\Omega) = \{v \in W^{1,p}(\Omega); \operatorname{tr} v = 0\}.$$

(d) 设 $1 \leqslant p < \infty$, 则定义为 $\mathcal{D}(\Omega)$ 在 $W^{2,p}(\Omega)$ 中闭包的空间 $W_0^{2,p}(\Omega)$ (6.5 节) 也可由下式给出[14)]

$$W_0^{2,p}(\Omega) = \left\{ v \in W^{2,p}(\Omega); \operatorname{tr} v = 0 \quad \text{且} \quad \sum_{i=1}^N \nu_i \operatorname{tr} \partial_i v = 0 \right\},$$

其中 $\nu = (\nu_i)_{i=1}^N$ 表示沿 Γ 的单位外法向量场 (其 $d\Gamma$ 几乎处处存在; 见 1.18 节). □

注意, 定理 6.6-5(a) 意味着

$$\operatorname{tr} \in \mathcal{L}(W^{1,p}(\Omega); L^p(\Gamma)), \quad p \geqslant 1,$$

总是成立的, 即与维数 N 无关.

当然, (c) 中的 tr 表示的线性连续算子, 它相应于 (a) 中所叙述的情形之一 (根据 p 与 N 大小的比较). 例如, 如果 $1 \leqslant p < N$, (c) 中的关系式 $\operatorname{tr} v = 0$ 就意味着函数 $\operatorname{tr} v$ 是空间 $L^{p^\#}(\Gamma)$ 中的零函数.

关系式 $\operatorname{tr}(W^{1,p}(\Omega)) \subsetneqq L^{p^\#}(\Gamma)$ 是定义下述迹空间的基础:

$$W^{1-\frac{1}{p},p}(\Gamma) := \{\operatorname{tr} v \in L^{p^\#}(\Gamma); v \in W^{1,p}(\Omega)\}, \quad \text{若 } 1 < p < N,$$
$$H^{1/2}(\Gamma) := \{\operatorname{tr} v \in L^2(\Omega); v \in H^1(\Omega)\}, \quad \text{若 } p = 2,$$

它们分别由 $W^{1,p}(\Omega)$, $1 \leqslant p < N$ 或 $H^1(\Omega)$ 中所有函数的迹组成.

根据惯常的做法, 此后在不至于引起混乱的情况下, 我们将省去符号 "tr". 例如, 定理 6.6-5 中的关系式 (c) 将简单地写为

$$W_0^{1,p}(\Omega) = \{v \in W^{1,p}(\Omega); v = 0 \text{ 在 } \Gamma \text{ 上}\} \quad \text{或}$$
$$H_0^1(\Omega) = \{v \in H^1(\Omega); v = 0 \text{ 在 } \Gamma \text{ 上}\}, \text{若 } p = 2.$$

---

14) 这个结果的证明, 参阅如 Nečas [1967, 第 2 章, 定理 4.12]. 类似的结果对 $m \geqslant 3$ 成立, 如果 Γ 是 $\mathcal{C}^{m,1}$ 类的; 参阅本著作中的定理 4.13.

我们也将碰到如下述的空间

$$V := \{v \in W^{1,p}(\Omega); \mathrm{tr}\, v = 0 \text{ 在 } \Gamma_0 \text{ 上}\},$$

其中 Γ_0 是 Γ 的 $d\Gamma$ 可测子集, 而关系式 "$\mathrm{tr}\, v = 0$ 在 Γ_0 上" 类似地意味着, 作为空间 $L^{p^\#}(\Gamma)$ 中的函数, $\mathrm{tr}\, v$ 在 Γ 的子集 Γ_0 上为零. 而同样地, 我们也将这样的一个空间 V 写为

$$V = \{v \in W^{1,p}(\Omega); v = 0 \text{ 在 } \Gamma_0 \text{ 上}\}.$$

对 $\mathbb{R}^N$ 中任意具有有限宽度的开子集成立的 Poincaré-Friedrichs 不等式 (定理 6.5-2), 当 Ω 是一区域时容许有下述有用的推广 (当 $p = 2$ 时, 结论 (a) 的证明是习题 6.7-7(1); 当 $p = 2$ 时 (b) 的证明将在定理 6.7-5 中给出).

定理 6.6-6 (广义的 Poincaré-Friedrichs 不等式) 设 Ω 为 $\mathbb{R}^N$ 中的区域, $1 \leqslant p < \infty$.

(a) 存在常数 c_0 使得

$$\int_\Omega |v|^p dx \leqslant c_0 \left\{ \int_\Omega \sum_{i=1}^N |\partial_i v|^p dx + \left| \int_\Omega v dx \right|^p \right\} \quad \text{对所有 } v \in W^{1,p}(\Omega).$$

(b) 设 Γ_0 是 Γ 的 $d\Gamma$ 可测子集, 且 $d\Gamma$-meas$\Gamma_0 > 0$. 则存在常数 c_2 使得

$$\|v\|_{1,p,\Omega} \leqslant c_2 |v|_{1,p,\Omega} \quad \text{对所有满足在 } \Gamma_0 \text{ 上 } v = 0 \text{ 的 } v \in W^{1,p}(\Omega).$$

(c) 设 Γ_0 是 Γ 的 $d\Gamma$ 可测子集, 且 $d\Gamma$-meas$\Gamma_0 > 0$. 则存在常数 c_1 使得

$$\int_\Omega |v|^p dx \leqslant c_1 \left\{ \int_\Omega \sum_{i=1}^N |\partial_i v|^p dx + \left| \int_{\Gamma_0} v d\Gamma \right|^p \right\} \quad \text{对所有 } v \in W^{1,p}(\Omega).$$

最后, 我们以将光滑函数的基本 Green 公式 (1.18 节) 的适用范围拓展到 Sobolev 空间中的函数, 来结束本节的讲述.

定理 6.6-7 (Sobolev 空间中的基本 Green 公式) 设 Ω 是 $\mathbb{R}^N$ 中的一区域, 而 $\nu = (\nu_i)_{i=1}^N$ 表示沿 Γ 的单位外法向量场. 设 $1 \leqslant p < \infty$ 及 $1 \leqslant q < \infty$ 使得

$$\begin{aligned} \frac{1}{p} + \frac{1}{q} \leqslant 1 + \frac{1}{N}, \quad & \text{若 } 1 \leqslant p < N \text{ 及 } 1 \leqslant q < N \\ & \text{或 } 1 < q, \text{ 若 } N \leqslant p \\ & \text{或 } 1 < p, \text{ 若 } N \leqslant q. \end{aligned}$$

则给定函数 $u \in W^{1,p}(\Omega)$ 及 $v \in W^{1,q}(\Omega)$, 每个函数 $uv\nu_i, 1 \leqslant i \leqslant N$, 都属于空间 $L^1(\Gamma)$ 而且

$$\int_\Omega u \partial_i v dx = -\int_\Omega (\partial_i u) v dx + \int_\Gamma uv\nu_i d\Gamma.$$

注意, 如果 $u, v \in H^1(\Omega)$, 则上述基本 Green 公式对任何维数 $N \geqslant 2$ 都成立.

关于 Sobolev 空间更专门的一些结果 (如其他 Green 公式或迹空间的各种稠密性定理等), 只要对具体的边值问题进行分析时需要, 我们还会予以陈述.

习题

6.6-1 这个习题的目的是证明, 当维数 $N \geqslant 2$ 时, 包含关系 $H^1(\Omega) \subset \mathcal{C}^0(\overline{\Omega})$ 不成立.

(1) 给定任意的 $0 < \rho < 1$, 设 $\Omega := \{x \in \mathbb{R}^2; |x| < \rho\}$. 证明, 对任何 $0 < \alpha < 1/2$, 在 Ω 中由 $u(x) := (-\ln|x|)^\alpha$ 若 $x \neq 0$, 几乎处处定义的函数 u 属于 Sobolev 空间 $H^1(\Omega)$ (如通常那样, 在函数与其等价类之间, 不加区分). 所以 $H^1(\Omega) \not\subset \mathcal{C}^0(\overline{\Omega})$.

(2) 设 $B := \bigcup_{k=1}^{\infty}\{b_k\}$ 是 Ω 的可列无限稠密子集, 而 $\beta_k > 0, k \geqslant 1$, 满足 $\displaystyle\sum_{k=1}^{\infty} \beta_k < \infty$. 证明, 在 Ω 中由 $v(x) := \displaystyle\sum_{k=1}^{\infty} \beta_k u(x - b_k)$ 若 $x \notin B$, 几乎处处定义的函数 v 属于 Sobolev 空间 $H^1(\Omega)$.

(3) 证明函数 v 到集合 Ω 上的任何延拓 $\widetilde{v}$ 在 Ω 中都是处处不连续的 (即不管指定 $\widetilde{v}(b_k)$, $k \geqslant 1$, 为何值). 注意, 任一这种延拓 $\widetilde{v}$ 也都属于 $H^1(\Omega)$, 此因 $\widetilde{v} = v$ 在 Ω 中几乎处处成立.

(4) 假定 $N \geqslant 3$ 而 $\Omega := \{x \in \mathbb{R}^N; |x| < 1\}$. 证明, 对任何 $0 < \lambda < (N-2)/2$, 在 Ω 中由 $w(x) := |x|^{-\lambda}$ 若 $x \neq 0$, 几乎处处定义的函数 w 属于空间 $H^1(\Omega)$. 因此 $H^1(\Omega) \not\subset \mathcal{C}^0(\overline{\Omega})$.

6.6-2 这个习题的目的是证明定理 6.6-1 在 $N = 1$ 这一特殊情况. 下面, $I :=]0, 1[$.

(1) 设 v 是 $H^1(I)$ 中的一个函数, 即 $v \in L^2(I)$ 且存在函数 $v_1 \in L^2(I)$ 使得 $\displaystyle\int_0^1 v\varphi' dx = -\int_0^1 v_1 \varphi dx$ 对所有 $\varphi \in \mathcal{D}(I)$ 成立. 证明由 $w(x) = \displaystyle\int_0^x v_1(t)dt, 0 \leqslant x \leqslant 1$, 定义的函数 $w : \overline{I} \to \mathbb{R}$ 属于空间 $\mathcal{C}(\overline{I})$ 且满足 $\displaystyle\int_0^1 (v - w)\varphi' dx = 0$ 对所有 $\varphi \in \mathcal{D}(I)$.

(2) 证明 $(v - w)$ 是一个常函数 (故由 (1), $v \in \mathcal{C}(\overline{I})$) 而且 $v(y) = v(x) + \displaystyle\int_x^y v_1(t)dt$ 对所有 $0 \leqslant x \leqslant y \leqslant 1$.

(3) 证明 $H^1(I) \hookrightarrow \mathcal{C}(\overline{I})$, 即 $H^1(I) \subset \mathcal{C}(\overline{I})$ 且存在一常数 c 使得

$$\sup_{0 \leqslant x \leqslant 1} |v(x)| \leqslant c\|v\|_{H^1(I)} \quad 对所有\ v \in H^1(I).$$

6.6-3 以下陈述是否正确? 设 Ω 为 $\mathbb{R}^3$ 的有界开子集, 其边界是有限个平面多边形的并. 则 Ω 是 $\mathbb{R}^3$ 中的一个区域.

6.6-4 (1) 证明 $\mathbb{R}^N$ 中一区域的边界的 N 维 Lebesgue 测度为零.

(2) 是否存在 $\mathbb{R}^N$ 的开子集, 其边界有大于零的 N 维 Lebesgue 测度?

6.6-5 设 Ω 是 $\mathbb{R}^N$ 中一区域, $1 \leqslant p \leqslant \infty$, 对每个整数 $m \geqslant 1$, 设 $\mathcal{P}_m(\Omega)$ 表示具有 N 个变量, 阶数 $\leqslant m$ 的所有多项式在 Ω 上的限制形成的空间.

(1) 证明, 半范数 $|\cdot|_{m+1,p,\Omega}$ 是商空间 $W^{m+1,p}(\Omega)/\mathcal{P}_m(\Omega)$ 上的一个范数, 它等价于这个空间上的商范数.

(2) 设 $\ell \in (W^{m+1,p}(\Omega))'$ 使得 $\ell(p) = 0$ 对所有 $p \in \mathcal{P}_m(\Omega)$. 证明存在一个与 ℓ 无关的常数 C 使得

$$|\ell(v)| \leqslant C\|\ell\|_{(W^{m+1,p}(\Omega))'}|v|_{m+1,p,\Omega} \quad 对所有\ v \in W^{m+1,p}(\Omega).$$

注 (2) 的结果构成 Bramble-Hilbert 引理[15].

6.7 二阶线性椭圆边值问题的例; 薄膜问题

现在, 必要的准备 (如二次极小问题或抽象变分问题, 以及 Sobolev 空间等) 已经就绪, 我们可以集中精力描述并分析提出于 $\mathbb{R}^N$ 中一区域上的各种边值问题实例. 在每一种情况, 我们都将遵循同样的三个步骤:

首先, 选定一个 Hilbert 空间 $V(H^1(\Omega)$ 或 $H^2(\Omega)$, 或者其闭子空间如 $H_0^1(\Omega)$ 或 $H_0^2(\Omega))$, V 的一个非空闭凸子集 U (在本节与下节中这种情况, 就可简单地取其等于 V), 一个双线性形式 $a: V \times V \to \mathbb{R}$ 以及一个线性形式 $\ell: V \to \mathbb{R}$.

其次, 如果 a 是对称的, 我们要验证所选定的 V, U, a 及 ℓ 满足定理 6.1-1 中关于存在性结果的所有假设. 如果验证下来确实如此, 则存在一个且只有一个函数 $u \in U$ 满足

$$J(u) = \inf_{v \in U} J(v), \quad 其中\ J(v) := \frac{1}{2}a(v,v) - \ell(v) \quad 对所有\ v \in V$$

或等价地, 满足变分方程

$$a(u,v) = \ell(v) \quad 对所有\ v \in U,$$

如果 $U = V$; 或满足变分不等式

$$a(u, v-u) \geqslant \ell(v-u) \quad 对所有\ v \in U,$$

如果 U 不是 V 的子空间 (定理 6.1-2). 若 a 不是对称的且 $U = V$, 则要借助于 Lax-Milgram 引理 (定理 6.2-1), 它断定满足下述变分方程的 $u \in V$ 的存在唯一性: $a(u,v) = \ell(v)$ 对所有 $v \in V$ (这种例子, 可见习题 6.7-9).

最后, 在关于函数 $u \in U$ 的附加的正则性假设下 (如 $u \in H^2(\Omega) \bigcap V$ 若 $V \subset H^1(\Omega)$ 或 $u \in H^4(\Omega) \bigcap V$ 若 $V \subset H^2(\Omega)$), 我们给出了上述变分方程或不等式的解 $u \in U$ 在 $\overline{\Omega}$ 上所满足的边值问题. 如果 $U = V$, 这个问题包括 u 在 Ω 中满足的形如 $\mathcal{L}u = f$ 的线性偏微分方程以及 u 在 Ω 的边界 Γ 上满足的线性边界条件. 术语 "边值问题" 反映出 u 在整个边界 Γ 上要满足条件. 注意, 边界条件的类型沿 Γ 是可以变化的 (这种例子, 可见定理 6.7-6).

[15] J.H. Bramble; S.R. Hilbert [1970]: Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation. SIAM Journal on Numerical Analysis **7**, 112–124.

在实例的选取方面, 我们按 “从最简单的到不太简单的” 进行. 这就是为什么在讨论更一般的二阶椭圆偏微分算子之前, 我们先从考虑形如 $\mathcal{L}: v \to \mathcal{L}v := -\Delta v + cv$ 的线性偏微分算子 $\mathcal{L}$ 开始; 这就是为什么我们顺序地先考虑 Dirichlet, 然后 Neumann, 最后是混合边界条件; 这就是为什么在讨论四阶偏微分算子 (6.8 节) 之前我们先考虑二阶的; 这也就是为什么在讨论非线性问题 (当 U 不是 V 的子空间时, 6.9 节) 之前我们先考虑线性问题 (当 $U = V$ 时).

为在本节中讨论的实例作准备, 我们证明两个有用的 Sobolev 空间中的 Green 公式. 回忆一下, 沿一个区域的边界 Γ, 单位外法向量场是 $d\Gamma$ 几乎处处存在的 (1.18 节).

任意足够光滑的函数: $u, v : \Omega \to \mathbb{R}$, 令

$$\nabla v := (\partial_i v)_{i=1}^N, \quad |\nabla v| := \left(\sum_{i=1}^N |\partial_i v|^2\right)^{1/2} \text{ 且 } \nabla u \cdot \nabla v := \sum_{i=1}^N \partial_i u \partial_i v.$$

回忆 $|a|$ 和 $a \cdot b$ 分别表示 $a \in \mathbb{R}^N$ 的范数和 $a, b \in \mathbb{R}^N$ 的 Euclid 内积.

定理 6.7-1　设 Ω 为 $\mathbb{R}^N$ 中的区域, $\nu = (\nu_i)_{i=1}^N$ 表示沿 $\Gamma := \partial\Omega$ 的单位外法向量场.

(a) 对任意 $u \in H^2(\Omega)$, 令

$$\Delta u := \sum_{i=1}^N \partial_{ii} u \in L^2(\Omega) \quad \text{及 } \partial_\nu u := \sum_{i=1}^N \nu_i \partial_i u \in L^2(\Gamma),$$

其中 $\partial_i u \in L^2(\Gamma)$ 表示函数 $\partial_i u \in H^1(\Omega)$ 在 Γ 上的迹. 则下述 Green 公式成立:

$$\int_\Omega \nabla u \cdot \nabla v dx = -\int_\Omega (\Delta u) v dx + \int_\Gamma (\partial_\nu u) v d\Gamma \quad \text{对所有 } u \in H^2(\Omega), v \in H^1(\Omega).$$

(b) 给定函数 $a \in \mathcal{C}^1(\overline{\Omega})$ 及 $u \in H^1(\Omega)$, 函数 au 属于空间 $H^1(\Omega)$. 此外, 对所有 $1 \leqslant j \leqslant N$ 下述 Green 公式成立:

$$\int_\Omega au \partial_j v dx = -\int_\Omega (\partial_j (au)) v dx + \int_\Gamma auv\nu_j d\Gamma \quad \text{对所有 } u \in H^1(\Omega), v \in H^1(\Omega).$$

证明　如果 $u \in H^2(\Omega)$, 则空间 $H^m(\Omega)$ 的定义 (6.5 节) 意味着每个函数 $\partial_i u, 1 \leqslant i \leqslant N$, 都属于空间 $H^1(\Omega)$. 既然如此, 第一个 Green 公式就简单地从 Sobolev 空间中的基本 Green 公式 (定理 6.6-7) 得到.

如果 $a \in \mathcal{C}^1(\overline{\Omega})$ 且 $u \in H^1(\Omega)$, 则函数 au 属于空间 $L^2(\Omega)$. 此外, 函数

$$w_j := (\partial_j a) u + a \partial_j u, \quad 1 \leqslant j \leqslant N,$$

显然属于 $L^2(\Omega)$ 而且是 au 的弱导数. 为说明这一点, 只需注意, 若 $u \in \mathcal{C}^\infty(\overline{\Omega})$, 则

$$\int_\Omega au \partial_j \varphi dx = -\int_\Omega \{(\partial_j a) u \varphi + a(\partial_j u)\varphi\} dx = -\int_\Omega w_j \varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega).$$

于是 $\mathcal{C}^\infty(\overline{\Omega})$ 在 $H^1(\Omega)$ 中的稠密性 (定理 6.6-4), 再结合 $L^2(\Omega)$ 中内积的连续性, 就证明了函数 $w_j \in L^2(\Omega), 1 \leqslant j \leqslant N$, 实际上正是 au 的弱导数.

所以函数 au 属于空间 $H^1(\Omega)$. 第二个 Green 公式又可简单地再次利用 Sobolev 空间中的基本 Green 公式得到. □

算子

$$\Delta := \sum_{i=1}^{N} \partial_{ii}$$

作用在定义于 Ω 内的函数上, 称为 Laplace 算子, 而 Δu 称为 u 的 Laplace 算符 (Laplacian). 算子

$$\partial_\nu := \sum_{i=1}^{N} \nu_i \partial_i$$

作用于定义在边界 Γ 上的函数上, 称为外法向微分算子, 而 $\partial_\nu u$ 称为 u 的外法向导数.

注 外法向微分算子在定理 6.6-5(d) 中已经遇到过.

我们现在讨论第一个实例.

定理 6.7-2 设 Ω 是 $\mathbb{R}^N$ 中具有有限宽度的开子集, 函数

$$c \in L^\infty(\Omega) \quad 且\ c \geqslant 0 \quad 在\ \Omega\ 中几乎处处成立,\ f \in L^2(\Omega)$$

是给定的, 又令

$$\begin{aligned}
V = U : &= H_0^1(\Omega), \\
a(u,v) : &= \int_\Omega (\nabla u \cdot \nabla v + cuv)dx \quad 对所有\ u,v \in V, \\
\ell(v) : &= \int_\Omega fv dx \quad 对所有\ v \in V.
\end{aligned}$$

则存在唯一的函数 $u \in H_0^1(\Omega)$, 它在空间 $H_0^1(\Omega)$ 上极小化由下式定义的泛函 $J : H_0^1(\Omega) \to \mathbb{R}$,

$$\begin{aligned}
J(v) : &= \frac{1}{2}a(v,v) - \ell(v) \\
&= \frac{1}{2}\int_\Omega (|\nabla v|^2 + cv^2)dx - \int_\Omega fv dx \quad 对所有\ v \in H_0^1(\Omega)
\end{aligned}$$

或等价地, 满足变分方程:

$$\int_\Omega (\nabla u \cdot \nabla v + cuv)dx = \int_\Omega fv dx \quad 对所有\ v \in H_0^1(\Omega).$$

此外, 以这种方式定义的线性映射

$$f \in L^2(\Omega) \to u \in H_0^1(\Omega)$$

是连续的.

最后, 函数 u 满足下述边值问题:

$$-\Delta u + cu = f \quad \text{在 } \Omega \text{ 中, 且 } u = 0 \text{ 在 } \Gamma \text{ 上}.$$

其中 Ω 中的偏微分方程理解为空间 $\mathcal{D}'(\Omega)$ 中的等式, 在 Ω 是区域的补充假设下, Γ 上的边界条件理解为空间 $L^2(\Gamma)$ 中的等式.

证明 对称的双线性形式 $a(\cdot,\cdot)$ 是连续的, 因为首先对 $L^2(\Omega)$ 中的函数, 然后对 $\mathbb{R}^{N+1}$ 中的向量利用 Cauchy-Schwarz 不等式, 有

$$\begin{aligned}|a(u,v)| &\leqslant \sum_{i=1}^{N} \|\partial_i u\|_{0,\Omega}\|\partial_i v\|_{0,\Omega} + \|c\|_{L^\infty(\Omega)}\|u\|_{0,\Omega}\|v\|_{0,\Omega} \\ &\leqslant \max\{1, \|c\|_{L^\infty(\Omega)}\}\left(\sum_{i=1}^{N}\|\partial_i u\|_{0,\Omega}^2 + \|u\|_{0,\Omega}^2\right)^{1/2}\left(\sum_{i=1}^{N}\|\partial_i v\|_{0,\Omega}^2 + \|v\|_{0,\Omega}^2\right)^{1/2} \\ &= \max\{1, \|c\|_{L^\infty(\Omega)}\}\|u\|_{1,\Omega}\|v\|_{1,\Omega} \quad \text{对所有 } u, v \in H^1(\Omega).\end{aligned}$$

另外, 双线性形式 a 是 $H_0^1(\Omega)$ 强制的, 因

$$a(v,v) \geqslant \int_\Omega |\nabla v|^2 dx = |v|_{1,\Omega}^2 \quad \text{对所有 } v \in H^1(\Omega),$$

而且半范数 $|\cdot|_{1,\Omega}$ 是空间 $H_0^1(\Omega)$ 上的一个范数, 等价于范数 $\|\cdot\|_{1,\Omega}$ (定理 6.5-2). 最后, 线性形式 ℓ 也是连续的, 因为

$$|\ell(v)| \leqslant \|f\|_{0,\Omega}\|v\|_{0,\Omega} \leqslant \|f\|_{0,\Omega}\|v\|_{1,\Omega} \quad \text{对所有 } v \in H^1(\Omega).$$

因此定理 6.1-1 中的所有假设都满足, 于是存在一个且只有一个函数 $u \in H_0^1(\Omega)$, 它在 $H_0^1(\Omega)$ 上极小化所宣示的泛函 J; 或等价地, 满足所示的变分方程 (定理 6.1-2).

映射 $f \in L^2(\Omega) \to u \in H_0^1(\Omega)$ 显然是线性的, 其连续性可由下述不等式得到. 解 u 满足

$$|u|_{1,\Omega}^2 \leqslant a(u,u) = \ell(u) \leqslant \|f\|_{0,\Omega}\|u\|_{0,\Omega} \leqslant \|f\|_{0,\Omega}\|u\|_{1,\Omega},$$

再利用所有函数 $v \in H_0^1(\Omega)$ 满足的不等式

$$\|v\|_{1,\Omega} \leqslant C(\Omega)|v|_{1,\Omega},$$

就给出

$$\|u\|_{1,\Omega} \leqslant C(\Omega)^2\|f\|_{0,\Omega} \quad \text{对所有 } f \in L^2(\Omega).$$

由于 $\int_\Omega \nabla u \cdot \nabla v dx = -\langle \Delta u, v\rangle$ 对所有 $v \in \mathcal{D}(\Omega)$, 其中 $\langle\cdot,\cdot\rangle :=_{\mathcal{D}'(\Omega)} \langle\cdot,\cdot\rangle_{\mathcal{D}(\Omega)}$ (6.3 节) 而 Δu 理解为一个分布, 这样方程 $a(u,v) = \ell(v)$ 对所有 $v \in V$ 就意味着

$$\langle -\Delta u + Cu - f, v\rangle = 0 \quad \text{对所有 } v \in \mathcal{D}(\Omega)$$

(此因 $\mathcal{D}(\Omega) \subset H_0^1(\Omega)$), 所以有

$$-\Delta u + Cu = f \quad \text{在 } \mathcal{D}'(\Omega) \text{ 中}.$$

当 Ω 是一个区域时, 关于空间 $H_0^1(\Omega)$ 的刻画 (定理 6.6-5(c)) 说明, 函数 $u \in H_0^1(\Omega)$ 在 Γ 上满足边界条件 $u = 0$, 在此这个边界条件被解释为在空间 $L^2(\Gamma)$ 中的等式. □

出现在定理 6.7-2 中的边值问题

$$-\Delta u + cu = f \quad \text{在 } \Omega \text{ 中}, u = 0 \text{ 在 } \Gamma \text{ 上}$$

或更一般地, 边值问题

$$-\Delta u + cu = f \text{ 在 } \Omega \text{ 中}, u = u_0 \text{ 在 } \Gamma \text{ 上}$$

(它同样可以从变分问题导出; 参阅习题 6.7-1) 称为*偏微分算子*

$$\mathcal{L} : v \to \mathcal{L}v := -\Delta v + cv$$

的 Dirichlet *问题*[16].

要注意, 在此及以后, 要么把出现在算子 $\mathcal{L}$ 定义中的函数 v 理解为至少在 Ω 中几乎处处有定义的, 而且足够光滑使算子定义有意义 (例如, $\mathcal{L}v \in L^2(\Omega)$ 若 $v \in H^2(\Omega)$); 或者把算子 $\mathcal{L}$ 理解为在*分布意义下*的线性偏微分算子 (6.3 节).

如果 $u_0 = 0$, 边界条件 $u = u_0$ 在 Γ 上称为*齐次的*, 否则称为*非齐次的* Dirichlet *边界条件*.

对于方程 $-\Delta u + cu = f$ 在 Ω 中, 其 $c = 0$ 及 $c = f = 0$ 的特殊情况, 即

$$-\Delta u = f \quad \text{在 } \Omega \text{ 中}, \ -\Delta u = 0 \text{ 在 } \Omega \text{ 中},$$

分别称为 Poisson *方程*[17] 和 Laplace *方程*[18]. 这些方程是非常重要的, 因为它们是范围极为广泛的各种物理现象的模型 (下面将给出一个相应于 $N = 2$ 的实例). 此外, 尽管形式异常简洁, 但 Laplace 方法还是一个完整的数学理论 —— *调和函数*理论的根基所在. 所谓调和函数, 就是在 Ω 中满足 $-\Delta u = 0$ 的那种函数 $u \in \mathcal{C}^2(\Omega)$ (在此, Ω 可以是 $\mathbb{R}^N$ 的任意开子集). 我们将调和函数的一些值得关注的性质简略地罗列于习题 6.7-3 到 6.7-6 中[19].

---

[16] 冠名源自 Gustav Lejeune Dirichlet (1805—1859).

[17] 冠名源自 Siméon–Denis Poisson (1781—1840).

[18] 冠名源自 Pierre–Simon Laplace (1749—1827).

[19] 关于调和函数理论富有启迪的介绍可参阅 Evans [2010, 2.2 节] Gilbarg & Trudinger [1998, 第 2 章].

对于第一个实例, 下面依序给出几个重要的注释. 第一, 它提供一个适定问题[20]的例子. 适定的含义是, 对任意的 $f \in L^2(\Omega)$ 存在唯一的解 $u \in H_0^1(\Omega)$, 此外该解还连续地依赖函数 $f \in L^2(\Omega)$.

注 在 7.10 节中我们将会看到, 利用极大值原理也可以确立解 u 对于右端项 f 的连续依赖性, 但在那里是关于 sup 范数的.

第二, 定理 6.7-2 证明中的最后一部分说明, 解释偏微分方程 $-\Delta u + cu = f$ 在 Ω 中的正确方法是视其为在空间 $\mathcal{D}'(\Omega)$ 中的等式. 但与之相比, 如果 Ω 是个区域, 边界条件 $u = 0$ 在 Γ 上总被视为空间 $L^2(\Gamma)$ 中函数的等式.

注 实际上, 定理 6.7-2 的所有结论在更一般的情况下, 即函数 $f \in L^2(\Omega)$ 换为分布 $f \in H^{-1}(\Omega)$ (空间 $H^{-1}(\Omega)$ 表示空间 $H_0^1(\Omega)$ 的对偶空间, 将在 6.11 节中定义), 仍然成立.

第三, 要确保 $u \in H^2(\Omega)$, 此时出现在 Δu 定义中的每个偏导数 $\partial_{ii}u \in \mathcal{D}'(\Omega)$ 都是 $L^2(\Omega)$ 中的函数, 其充分条件是一个很精巧的问题[21]. 例如, 可以证明, 如果边界 Γ 是 $\mathcal{C}^2$ 类的 (1.18 节), 对任意函数 $f \in L^2(\Omega)$, 变分方程

$$\int_\Omega (\nabla u \cdot \nabla v + uv)dx = \int_\Omega fvdx \quad \text{对所有} v \in H_0^1(\Omega)$$

的解 $v \in H_0^1(\Omega)$ 属于空间 $H^2(\Omega)$[22] (在这方面, 习题 6.7-2 可作为有趣的补充).

第四, 要保证 u 是这样一个边值问题的经典解, 即 u 在空间 $\mathcal{C}(\overline{\Omega}) \bigcap \mathcal{C}^2(\Omega)$ 内, 需要数据资料的进一步的假设.

在这个方向上, 当 $c = f = 0$ 时, 一个完美的定理[23] 断言, 如果开集 $\Omega \subset \mathbb{R}^N$ 是有界的, Dirichlet 问题: $-\Delta u = 0$ 在 Ω 中及 $u = u_0$ 在 Γ 上, 对任意函数 $u_0 \in \mathcal{C}(\Gamma)$ 有 (唯一) 解 $u \in \mathcal{C}(\overline{\Omega}) \bigcap \mathcal{C}^2(\Omega)$ 的充分必要条件是, 给定任一点 $y \in \Gamma$, 存在一个闸函数, 即一个函数 $w_y \in \mathcal{C}(\overline{\Omega}) \bigcap \mathcal{C}^2(\Omega)$ 具有下述性质: $-\Delta w_y \geqslant 0$ 在 Ω 中, $w_y(y) = 0$ 及 $w_y(x) > 0$ 对所有 $x \in (\overline{\Omega} - \{y\})$ (例如, 如果边界 Γ 是 $\mathcal{C}^2$ 类的就属于这种情况; 但

---

[20] 适定问题的概念是 Jacques Hadamard (1865—1963) 在下文中引入的:

J. Hadamard [1902]: Sur les problèmes aux dérivées partielles et leur signification physique. Princeton University Bulletin **13**, 49–52.

Acaptivating account of the adventurous life of Jacques Hadamard is given in Maz'ya & Shaposhnikova [1998].

[21] 如果 Ω 只是一个区域, 满足 $\Delta u \in L^2(\Omega)$ 的函数 $u \in H_0'(\Omega)$ 不一定属于 $H^2(\Omega)$; 反例可见:

D. Jerison; C. E. Kenig [1995]: The inhomogeneous Dirichlet problem in Lipschitz domains. Journal of Functional Analysis **130**, 161–219.

[22] 例如, 可参阅 Brezis [2001, 定理 9.25] 或 Evans [2010, 6.3 节].

[23] 结果属于:

O. Perron [1923]: Eine neue Behandlung der Randwertaufgabe für $\Delta u = 0$. Mathematische Zeitschrift **18**, 42–54.

Perron 方法的现代处理, 见 Gilbarg & Trudinger [1998, 第 2 章].

如果 Γ 只是 Lipschitz 连续的就不一定属于这种情况).

在一般情况, 为了得到经典解, 合适的配置是空间 $\mathcal{C}^{m,\lambda}(\overline{\Omega})$ (1.18 节) 中的函数. 对这种情况, 另一个完美的定理[24] 断言, 如果对某个 $0 < \lambda < 1$, 函数 $c \geqslant 0$ 属于空间 $\mathcal{C}^{0,\lambda}(\overline{\Omega})$ 且区域 Ω 的边界 Γ 是 $\mathcal{C}^{2,\lambda}$ 类的, 则对于任意函数 $f \in \mathcal{C}^{0,\lambda}(\Omega)$ 和 $u_0 \in \mathcal{C}^{2,\lambda}(\overline{\Omega})$, Dirichlet 问题: $-\Delta u + cu = f$ 在 Ω 中及 $u = u_0$ 在 Γ 上, 有 (唯一) 解 $u \in \mathcal{C}^{2,\lambda}(\overline{\Omega})$ (实际上, 这个存在结果对本节稍后定义的, 系数属于空间 $\mathcal{C}^{0,\lambda}(\overline{\Omega})$ 的一般椭圆型算子也成立). 这个定理依赖重要的 Schauder 估计[25], 该估计给出这个问题的范数 $\|u\|_{\mathcal{C}^{2,\lambda}(\overline{\Omega})}$ 一个先验的界.

当 $c = 0, N = 2$, 而 Ω 位于 "水平的" 平面内时, 这个问题是薄膜问题的数学模型, 它产生于如下的线性化弹性问题: 求一个张力为 τ 的弹性薄膜在密度为 $F = \tau f$ 的垂直力作用下的平衡位置, 使其垂直位移 $u : \overline{\Omega} \to \mathbb{R}$ 沿边界 Γ 等于已知函数 u_0 (图 6.7-1). 当 $u_0 = 0$ 时, 根据定理 6.7-2, 位移 $u \in H_0^1(\Omega)$ 在空间 $H_0^1(\Omega)$ 上极小化由下式定义的薄膜能量 $J : H_0^1(\Omega) \to \mathbb{R}$,

$$J(v) := \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \int_\Omega fv dx \quad \text{对所有 } v \in H_0^1(\Omega).$$

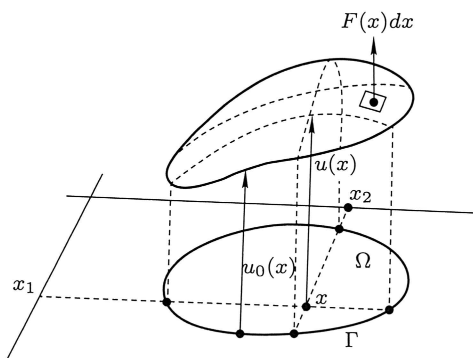


图 6.7-1 薄膜问题: 未知函数 $u : \overline{\Omega} \subset \mathbb{R}^2 \to \mathbb{R}$ 表示膜在单位面积密度为 F 的垂直力作用下的垂直位移. 该图出自于 P. G. Ciarlet [1978]: The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam

注 有趣的是, 出现于薄膜问题的变分形式中的资料, 可以通过将渐近分析方法用于非线性三维弹性方程组而得到严格的验证 (首先, 令厚度趋向零[26], 然后再令张

---

[24] 结果属于:

O. D. Kellogg [1929]: Foundations of Potential Theory. Springer, Berlin.

也见 Gilbarg & Trudinger [1998, 定理 6.14].

[25] J. Schauder [1934]: Über lineare elliptische Differentialgleichungen zweiter Ordnung. Mathematische Zeitschrift **38**, 257–282.

[26] P. G. Ciarlet [1980]: A justification of the von Kármán equations. Archive for Rational Mechanics and Analysis **73**, 349–389.

G. Friesecke; R. D. James; S. Müller [2006]: A hierarchy of plate models derived from nonlinear elasticity by Gamma–convergence. Archive for Rational Mechanics and Analysis **180**, 183–236.

力 τ 趋向 ∞[27]).

作为下一个实例的准备, 我们叙述一个关于迹的结果[28], 它对于导出在一个区域边界 Γ (定理 6.7-4) 或在 Γ 的子集 (定理 6.7-6) 上满足形如 $\partial_\nu u = 0$ 的边界条件有实质性作用.

定理 6.7-3 设 Ω 为 $\mathbb{R}^N$ 中的区域, Γ_1 是 $\Gamma := \partial\Omega$ 的相对开子集, 又设 $w \in L^2(\Gamma_1)$, 它还满足

$$\int_{\Gamma_1} wv d\Gamma = 0 \quad \text{对所有在 } \Gamma - \Gamma_1 \text{ 上满足 } v = 0 \text{ 的 } v \in H^1(\Omega).$$

则 $w = 0$.

注 根据定理 4.3-2, 定理 6.7-3 等价于空间 $\{v|_{\Gamma_1} \in L^2(\Gamma_1); v \in H^1(\Omega), v = 0$ 在 $\Gamma - \Gamma_1$ 上$\}$ 在 $L^2(\Gamma_1)$ 中稠密.

现在讨论我们的第二个实例. 与第一个例子相比, 它有以下几个不同之处: 开集 Ω 现在是一个区域, 更大的空间 V, 对于函数 c 的更强假设, 及更一般的线性形式 ℓ. 为简单起见, 下面证明中类似于定理 6.7-2 的证明中用过的那种推导就省略了.

定理 6.7-4 设 Ω 为 $\mathbb{R}^N$ 中的区域, 设函数

$$c \in L^\infty(\Omega) \text{ 且 } c \geqslant c_0 > 0 \quad \text{几乎处处在 } \Omega \text{ 中成立}, \ f \in L^2(\Omega), g \in L^2(\Gamma)$$

是给定的, 又设

$$\begin{aligned} V = U &:= H^1(\Omega), \\ a(u,v) &:= \int_\Omega (\nabla u \cdot \nabla v + cuv)dx \quad \text{对所有 } u, v \in V, \\ \ell(v) &:= \int_\Omega fvdx + \int_\Gamma gvd\Gamma \quad \text{对所有 } v \in V. \end{aligned}$$

则存在唯一的函数 $u \in H^1(\Omega)$, 它极小化由下式定义的泛函 $J : H^1(\Omega) \to \mathbb{R}$,

$$\begin{aligned} J(v) : &= \frac{1}{2}a(v,v) - \ell(v) \\ &= \frac{1}{2}\int_\Omega (|\nabla v|^2 + cv^2)dx - \left(\int_\Omega fvdx + \int_\Gamma gvd\Gamma\right) \end{aligned}$$

对所有 $v \in H^1(\Omega)$, 或等价地满足变分方程

$$\int_\Omega (\nabla u \cdot \nabla v + cuv)dx = \int_\Omega fvdx + \int_\Gamma gvd\Gamma \quad \text{对所有 } v \in H^1(\Omega).$$

---

[27] 见 Ciarlet & Rabier [1980, 2.3 节] 或 Ciarlet [1997, 5.10 节].

[28] 这个结果的证明, 在关于 Sobolev 空间的经典教科书中不常见, 可参阅:

J. M. E. Bernard [2011]: Density results in Sobolev spaces whose elements vanish on a part of the boundary. Chinese Annals of Mathematics, Series B, **32**, 823–846.

此外, 以这种方式定义的映射

$$(f, g) \in L^2(\Omega) \times L^2(\Gamma) \to u \in H^1(\Omega)$$

是连续的.

若进一步假设 $u \in H^2(\Omega)$, 则 u 满足下述边值问题:

$$-\Delta u + cu = f \quad \text{在 } \Omega \text{ 中}, \quad \partial_\nu u = g \quad \text{在 } \Gamma \text{ 上},$$

其中 $\partial_\nu u \in L^2(\Gamma)$ 表示 u 的外法向导数 (定理 6.7-1(a)).

证明　双线性形式是 $H^1(\Omega)$ 强制的, 此因

$$a(v, v) \geqslant \min\{1, c_0\}\|v\|_{1,\Omega}^2 \quad \text{对所有 } v \in H^1(\Omega).$$

线性形式 $v \in H^1(\Omega) \to \int_\Gamma gv d\Gamma$ 是连续的, 此因 (定理 6.6-5)

$$\begin{aligned}\left|\int_\Gamma gv d\Gamma\right| &\leqslant \|g\|_{L^2(\Gamma)}\|v\|_{L^2(\Gamma)} \\ &\leqslant \|\mathrm{tr}\|_{\mathcal{L}(H^1(\Omega);L^2(\Omega))}\|g\|_{L^2(\Gamma)}\|v\|_{1,\Omega}.\end{aligned}$$

所以存在唯一的函数 $u \in H^1(\Omega)$, 它在空间 $H^1(\Omega)$ 上极小化所示泛函 J 并且满足所示变分方程.

下面假设 $u \in H^2(\Omega)$. 利用定理 6.7-1(a) 中的 Green 公式, 方程 $a(u, v) = \ell(v)$ 对所有 $v \in V$, 变为

$$\int_\Omega (-\Delta u + cu - f)v dx = \int_\Gamma (g - \partial_\nu u)v d\Gamma \quad \text{对所有 } H^1(\Omega).$$

特别地, 函数 $(-\Delta u + cu - f) \in L^2(\Omega)$ 满足

$$\int_\Omega (-\Delta u + cu - f)v dx = 0 \quad \text{对所有 } v \in \mathcal{D}(\Omega).$$

因 $\mathcal{D}(\Omega)$ 在 $L^2(\Omega)$ 中稠密, 这意味着 $-\Delta u + cu - f = 0$ 在 $L^2(\Omega)$ 中成立.

考虑到 $-\Delta u + cu - f = 0$, 于是剩下来就有

$$\int_\Omega (g - \partial_\nu u)v d\Gamma = 0 \quad \text{对所有 } v \in H^1(\Omega),$$

这意味着函数 $(g - \partial_\nu u) \in L^2(\Gamma)$ 为零 (应用定理 6.7-3, 取 $\Gamma_1 = \Gamma$). □

出现在定理 6.7-4 中的边值问题

$$-\Delta u + cu = f \quad \text{在 } \Omega \text{ 中}, \quad \partial_\nu u = g \text{ 在 } \Gamma \text{ 上}$$

称为偏微分算子

$$\mathcal{L}: v \to \mathcal{L}v := -\Delta v + cv$$

的 Neumann[29] 问题. 边界条件 $\partial_\nu u = g$ 在 Γ 上, 如果 $g = 0$, 称为齐次的, 否则称为非齐次的 Neumann 边界条件.

要注意, 如果 u 只是属于 $H^1(\Omega)$, 即没有关于函数 u 的附加正则性假定, 出现在第一个实例中的 Dirichlet 边界条件 $u = 0$ 在 Γ 上, 在空间 $L^2(\Gamma)$ 中是有意义的; 但在这种情况下, Neumann 边界条件 $\partial_\nu u = g$ 在 Γ 上, 至少在任何 Γ 上的函数空间里没有意义. 然而, 即使 u 只属于 $H^1(\Omega)$, 我们仍将说, 至少在形式上, u 是边值问题: $-\Delta u + cu = f$ 在 Ω 中且 $\partial_\nu u = g$ 在 Γ 上的解.

注　定义空间

$$H(\Delta; \Omega) := \{v \in H^1(\Omega); \Delta v \in L^2(\Omega)\},$$

其中 $\Delta v \in L^2(\Omega)$ 理解为在分布意义下. 则可以证明[30] 存在一个连续线性算子 $\gamma_1 : H(\Delta; \Omega) \to H^{-1/2}(\Gamma)$, 其中 $H^{-1/2}(\Gamma)$ 表示迹空间 $H^{1/2}(\Gamma) := \{\mathrm{tr} v \in L^2(\Gamma); v \in H^1(\Omega)\}$ (6.6 节) 的对偶, 使得对充分光滑的函数 $v : \overline{\Omega} \to \mathbb{R}$ 成立 $\gamma_1 v = \partial_\nu v|_\Gamma$ (这里空间 $H(\Delta; \Omega)$ 装备以自然范数 $v \to (\|v\|_{1,\Omega}^2 + \|\Delta v\|_{0,\Omega}^2)^{1/2}$).

这一考虑给边界条件 $\partial_\nu u = g$ 在 Γ 上, 提供了一个确定的意义, 即可视其为对偶空间 $H^{-1/2}(\Gamma)$ 中的等式, 即使定理 6.7-3 中讨论的变分问题的解 u 只属于 $H^1(\Omega)$.

如果 Γ_1 是 Γ 的相对开的真子集, 形式边界条件 $\partial_\nu u = g$ 在 Γ_1 上 (这种边界条件在下一实例中出现) 的解释就更复杂, 即要视为空间

$$H_{00}^{1/2}(\Gamma_1) := \{v \in L^2(\Gamma_1);\quad \text{存在 } w \in H^1(\Omega),\quad w = 0 \text{ 在 } \Gamma - \Gamma_1 \text{ 上及 } w = v \text{ 在 } \Gamma_1 \text{上}\}$$

的对偶中的等式, 上述空间并不等于空间 $H^{1/2}(\Gamma_1) := \{v|_{\Gamma_1}; v \in H^{1/2}(\Gamma)\}$.

也要注意, 出现在第一个实例中的边界条件 $u = 0$ 在 Γ 上, 就简单地得自 u 所属空间 V 的定义, 即 $V = H_0^1(\Omega)$; 但出现在第二个实例中的边界条件 $\partial_\nu u = g$ 在 Γ 上, 不是直接来自函数空间而是得自 Green 公式的应用.

如果将假设 $c \geqslant c_0 > 0$ 在 Ω 中几乎处处成立换成 $c = 0$ 在 Ω 中几乎处处成立 (第一个实例的特殊情况), 则双线性形式 a 不再是 $H^1(\Omega)$ 椭圆的. 然而, 一个关于存在性的结果仍然成立, 只是要求函数 f 与 g 满足一个适当的相容性条件; 见习题 6.7-7 与 6.7-8.

回忆一下, 当 Ω 具有有限宽度时, 存在常数 $C = C(\Omega)$ 使得 Poincaré–Friedrichs 不等式成立, 即

$$\|v\|_{1,\Omega} \leqslant C|v|_{1,\Omega} \quad \text{对所有 } v \in H_0^1(\Omega)$$

(定理 6.5-2). 我们现在证明, 如果 Ω 是区域, 而出现在不等式中的函数 $v \in H^1(\Omega)$ 只在边界的部分 Γ_0 上为零, 假如 $d\Gamma$ 测度 $\Gamma_0 > 0$, 则以上不等式仍然成立. 下面的结果

---

[29] 冠名源自 Carl Neumann (1832—1925).

[30] 参阅, 如 DAUTRAY & LIONS [2000b, 第 7 章, 第 1 节].

在确立第三个实例中的双线性形式的椭圆性时需要, 它也附带地为定理 6.5-2(b) 关于区域当 $m=1$ 时的情况提供了另一个证明.

定理 6.7-5 设 Ω 为 $\mathbb{R}^N$ 中的区域, Γ_0 是边界 Γ 的 $d\Gamma$ 可测子集并满足

$$d\Gamma \text{ 测度 } \Gamma_0 > 0.$$

则空间

$$V := \{v \in H^1(\Omega); v = 0 \text{ 在 } \Gamma_0 \text{ 上}\}$$

是 $H^1(\Omega)$ 的闭子空间, 且存在常数 $C = C(\Omega)$ 使得

$$|v|_{1,\Omega} \leqslant \|v\|_{1,\Omega} \leqslant C|v|_{1,\Omega} \quad \text{对所有 } v \in V.$$

证明 设 $(v_k)_{k=1}^{\infty}$ 是空间 V 中收敛于函数 $v \in H^1(\Omega)$ 的函数序列. 因为此时序列 $(\text{tr } v_k)_{k=1}^{\infty}$ 在空间 $L^2(\Gamma)$ 中收敛于 $\text{tr } v$ (定理 6.6-5), 故 $(\text{tr } v_k|_{\Gamma_0})_{k=1}^{\infty}$ 在 $L^2(\Gamma_0)$ 中收敛于 $\text{tr } v|_{\Gamma_0}$. 但 $\text{tr } v_k|_{\Gamma_0} = 0$ 对所有 $k \geqslant 1$ 而且赋范向量空间中序列的极限是唯一的, 所以 $\text{tr } v|_{\Gamma_0} = 0$, 而 V 是 $H^1(\Omega)$ 的闭子空间.

其次, 我们证明 $|\cdot|_{1,\Omega}$ 是空间 V 上的一个范数. 设 v 是空间 V 中满足 $|v|_{1,\Omega} = 0$ 的一个函数. 则由于集合 Ω 的连通性, v 是常函数 (定理 6.3-4). 所以其在 Γ 上的迹是取同样值的常函数 (由迹算子的构造知, 对 $\mathcal{C}^{\infty}(\overline{\Omega})$ 中的函数, tr 与通常的迹是等同的; 见 6.6 节). 但因为在 $d\Gamma$ 测度 > 0 的集合 Γ_0 上, 迹为零, 故这个常数恒为零.

最后假设两个范数 $|\cdot|_{1,\Omega}$ 与 $\|\cdot\|_{1,\Omega}$ 在空间 V 上不等价. 则一定存在函数 $v_k \in V$ 的序列 $(v_k)_{k=1}^{\infty}$, 满足

$$\|v_k\|_{1,\Omega} = 1 \quad \text{对所有 } k \quad \text{且} \quad \lim_{k\to\infty} |v_k|_{1,\Omega} = 0.$$

由 Rellich-Kondrachov 定理 (定理 6.6-3), 空间 $H^1(\Omega)$ 中任何有界序列包含一个在空间 $L^2(\Omega)$ 中收敛的子列. 所以存在序列 $(v_k)_{k=1}^{\infty}$ 的子列 $(v_\ell)_{\ell=1}^{\infty}$, 它在空间 $L^2(\Omega)$ 中收敛.

另一方面, 因为 $\lim_{\ell\to\infty} |v_\ell|_{1,\Omega} = 0$, 因此序列 $(v_\ell)_{\ell=1}^{\infty}$ 是空间 V 中的 Cauchy 序列, 而作为 $H^1(\Omega)$ 的闭子空间 V 是完备的. 所以序列 $(v_\ell)_{\ell=1}^{\infty}$ 关于范数 $\|\cdot\|_{1,\Omega}$ 收敛于 $v \in V$.

因为 $|v|_{1,\Omega} = \lim_{\ell\to\infty} |v_\ell|_{1,\Omega} = 0$ 且 $v \in V$, 故 $v = 0$, 这与等式 $\|v_\ell\|_{1,\Omega} = 1$ 对所有 ℓ, 矛盾. 证明完成. □

在下面第三个实例中, 我们将在两个方向上拓展前面的实例: 首先, 两种类型的边界条件, 即 Dirichlet 与 Neumann 型边界条件都将出现在相应的边值问题中; 其次, 偏微分算子也将更一般.

定理 6.7-6 设 Ω 是 $\mathbb{R}^N$ 中的区域, $a_{ij} = a_{ji} \in \mathcal{C}^1(\overline{\Omega}), 1 \leqslant i, j \leqslant N$, 是给定的具

有以下性质的函数: 存在常数 μ 使得

$$\mu > 0 \quad \text{且} \quad \sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geqslant \mu \sum_{i=1}^{N} |\xi_i|^2 \quad \text{对所有 } x \in \overline{\Omega} \text{ 及所有 } (\xi_i)_{i=1}^{N} \in \mathbb{R}^N.$$

设 Γ_1 是 $\Gamma := \partial\Omega$ 的相对开子集使得

$$d\Gamma \text{ 测度 } \Gamma_0 > 0, \quad \text{其中 } \Gamma_0 = \Gamma - F_1,$$

又设函数

$$c \in L^\infty(\Omega) \text{ 使得 } c \geqslant 0 \text{ 在 } \Omega \text{ 中几乎处处成立}, f \in L^2(\Omega), g \in L^2(\Gamma_1)$$

是给定的. 最后, 设

$$\begin{aligned}
V = U : &= \{v \in H^1(\Omega); v = 0 \quad \text{在 } \Gamma_0 \text{ 上}\}, \\
a(u,v) : &= \int_\Omega \left(\sum_{i,j=1}^{N} a_{ij}\partial_i u \partial_j v + cuv \right) dx \quad \text{对所有 } u, v \in V, \\
\ell(v) : &= \int_\Omega fv dx + \int_{\Gamma_1} gv d\Gamma \quad \text{对所有 } v \in V.
\end{aligned}$$

则存在唯一的函数 $u \in V$ 在空间 V 上极小化由下式定义的泛函 $J : V \to \mathbb{R}$,

$$\begin{aligned}
J(v) : &= \frac{1}{2} a(v,v) - \ell(v) \\
&= \frac{1}{2} \int_\Omega \left(\sum_{i,j=1}^{N} a_{ij}\partial_i v \partial_j v + cv^2 \right) dx - \left(\int_\Omega fv dx + \int_{\Gamma_1} gv d\Gamma \right)
\end{aligned}$$

对所有 $v \in V$; 或等价地, 满足变分方程

$$\int_\Omega \left(\sum_{i,j=1}^{N} a_{ij}\partial_i v \partial_j v + cuv \right) dx = \int_\Omega fv dx + \int_{\Gamma_1} gv d\Gamma \quad \text{对所有 } v \in V.$$

此外, 以这种方式定义的线性映射

$$(f,g) \in L^2(\Omega) \times L^2(\Gamma_1) \to u \in V \subset H^1(\Omega)$$

是连续的.

进一步假定 $u \in H^2(\Omega)$. 则 u 满足下述边值问题

$$\begin{aligned}
-\sum_{i,j=1}^{N} \partial_j(a_{ij}\partial_i u) + cu = f \quad &\text{在 } \Omega \text{ 中}, \\
u = 0 \quad &\text{在 } \Gamma_0 \text{ 上}, \\
\sum_{i,j=1}^{N} a_{ij}\nu_j\partial_i u = g \quad &\text{在 } \Gamma_1 \text{ 上},
\end{aligned}$$

其中 $\nu_j, 1 \leqslant j \leqslant N$, 表示沿 Γ 单位外法向量的分量.

证明　作为 $H^1(\Omega)$ 的闭子空间, 空间 V 是 Hilbert 空间 (定理 6.7-5). 双线性形式 a 是连续的, 此因

$$|a(u,v)| \leqslant \max\{\|c\|_{L^\infty(\Omega)}, \max_{1\leqslant i,j\leqslant N} \|a_{ij}\|_{\mathcal{C}(\overline{\Omega})}\}\|u\|_{1,\Omega}\|v\|_{1,\Omega} \quad \text{对所有 } u,v \in H^1(\Omega).$$

双线性形式 a 也是 V 强制的, 此因

$$a(v,v) = \int_\Omega \left(\sum_{i,j=1}^N a_{ij}\partial_i v \partial_j v + cv^2\right) dx \geqslant \mu|v|_{1,\Omega}^2 \quad \text{对所有 } v \in V,$$

而且 V 中的范数 $|\cdot|_{1,\Omega}$ 等价于 $\|\cdot\|_{1,\Omega}$ (定理 6.7-5). 线性形式显然也是连续的.

所以存在唯一的函数 u 在空间 V 上极小化所示泛函 J, 或等价地, 满足所示变分方程.

下面假定 $u \in H^2(\Omega)$. 将定理 6.7-1(b) 中的 Green 公式应用于函数 $\partial_i u \in H^1(\Omega)$ (由假设, 函数 a_{ij} 属于空间 $\mathcal{C}^1(\overline{\Omega})$) 并注意关系式 $v = 0$ 在 Γ_0 上, 变分方程 $a(u,v) = \ell(v)$ 对所有 $v \in V$, 就变为

$$\int_\Omega \left(-\sum_{i,j=1}^N \partial_j(a_{ij}\partial_i u) + cu - f\right) v dx = \int_{\Gamma_1} \left(g - \sum_{i,j=1}^N a_{ij}\nu_j\partial_i u\right) v d\Gamma \quad \text{对所有 } v \in V.$$

特别地, 有

$$\int_\Omega \left(-\sum_{i,j=1}^N \partial_j(a_{ij}\partial_i u) + cu - f\right) v dx = 0 \quad \text{对所有 } v \in \mathcal{D}(\Omega),$$

这意味着 $(-\sum_{i,j=1}^N \partial_j(a_{ij}\partial_i u) + cu - f) = 0$ 在 $L^2(\Omega)$ 中成立. 考虑到这个方程, 剩下来的就是

$$\int_{\Gamma_1} \left(g - \sum_{i,j=1}^N a_{ij}\nu_j\partial_i u\right) v d\Gamma = 0 \quad \text{对所有 } v \in V,$$

由定理 6.7-3, 这意味着 $g - \sum_{i,j=1}^N a_{ij}\nu_j\partial_i u = 0$ 在 $L^2(\Gamma_1)$ 中成立. 最后, $u \in V$ 就意味着 $u = 0$ 在 Γ_0 上. □

定理 6.7-6 的证明阐明了要从变分方程复原二阶边值问题所需的三个步骤.

第一, 应用特定的 Green 公式 (假定变分方程解的充分正则性) 并令变分方程中的函数 v 在空间 $\mathcal{D}(\Omega)$ 中变动 (空间 V 总是包含空间 $\mathcal{D}(\Omega)$ 的). 这给出了偏微分方程 $\mathcal{L}u = f$, 它至少在分布意义下[31], 即在空间 $\mathcal{D}'(\Omega)$ 中成立.

第二, 考虑到方程 $\mathcal{L}u = f$ 在 Ω 内已经满足, 令 v 在整个空间 V 中变动. 剩下的变分方程只包含在 Γ_1 上的积分, 给出在 Γ_1 上的 Neumann 边界条件 (当然除了 $V = H_0^1(\Omega)$, 在这种情况下, 这一步就没必要了).

---

31) 对于在分布意义下偏微分方程特具启发性的讨论, 以及丰富的物理实例, 可参阅 SCHWARTZ [1965].

第三, 在 Γ_0 上的齐次 Dirichlet 边界条件可由 u 所属的空间 V 的定义得到, 从而完成了边值问题 (当然除了 $V = H^1(\Omega)$, 在这种情况下, 这一步就没必要了).

也要注意, 像本节中讨论的所有其他变分问题一样, 定理 6.7-6 中的变分问题在下述意义上是适定的, 即映射 $(f,g) \in L^2(\Omega) \times L^2(\Gamma_1) \to u \in V$ 是有确切定义且连续的.

为了反映其兼有 Dirichlet 与 Neumann 边界条件这一事实, 出现在定理 6.7-6 中的边值问题

$$
\begin{aligned}
-\sum_{i,j=1}^{N} \partial_j(a_{ij}\partial_i u) + cu = f \quad & \text{在 } \Omega \text{ 中}, \\
u = 0 \quad & \text{在 } \Gamma_0 \text{ 上}, \\
\sum_{i,j=1}^{N} a_{ij}\nu_j\partial_i u = g \quad & \text{在 } \Gamma_1 \text{ 上},
\end{aligned}
$$

称为偏微分算子 $\mathcal{L} : v \longrightarrow \mathcal{L}v := \sum_{i,j=1}^{N} \partial_j(a_{ij}\partial_i v) + cv$ 的混合问题.

出现在沿 Γ_1 的边界条件中的边界算子

$$
v \longrightarrow \sum_{i,j=1}^{N} a_{ij}\nu_j\partial_i v
$$

称为与偏微分算子 $\mathcal{L}$ 有关的余法向微分算子. 注意, 当 $\mathcal{L}v := -\Delta v + cv$ 时, 它就化为法向微分算子 ∂_ν.

下面, 以几个重要的定义来结束这一节.

所谓二阶线性偏微分算子 $\mathcal{L}$, 其对所有函数 $v \in \mathcal{C}^2(\Omega)$ 由形式如下的表达式给出

$$
\mathcal{L}v(x) = -\sum_{i,j=1}^{N} a_{ij}(x)\partial_{ij}v(x) + \sum_{i=1}^{N} b_i(x)\partial_i v(x) + c(x)v(x) \quad \text{对所有 } x \in \Omega,
$$

其中函数 $a_{ij}, b_i, c : \Omega \longrightarrow \mathbb{R}$ 为特定的系数. 如果矩阵 $(a_{ij}(x)), x \in \Omega$, 不失一般性总可假定其为对称的, 在每个 $x \in \Omega$ 都是正定的; 等价地, 即在每个 $x \in \Omega$, 存在一个常数 $\mu(x)$ 使得

$$
\mu(x) > 0 \text{ 且 } \sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geqslant \mu(x)\sum_{i=1}^{N} |\xi_i|^2 \quad \text{对所有 } (\xi_i)_{i=1}^{N} \in \mathbb{R}^N,
$$

则称该算子是椭圆的.

线性偏微分算子 $\mathcal{L}$ 称为一致椭圆的, 如果矩阵 $(a_{ij}(x)), x \in \Omega$, 在下述意义上是一致正定的: 存在常数 μ 使得

$$
\mu > 0 \text{ 且 } \sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geqslant \mu\sum_{i=1}^{N} |\xi_i|^2 \quad \text{对所有 } x \in \Omega \text{ 及所有 } (\xi_i)_{i=1}^{N} \in \mathbb{R}^N.
$$

如果出现在偏微分方程 $\mathcal{L}u = f$ (在 Ω 内) 的算子 $\mathcal{L}$ 是椭圆的或一致椭圆的二阶线性偏微分算子, 该线性边值问题称为二阶椭圆边值问题.

出现在本节所讨论的各个实例中的边值问题为我们提供一些二阶椭圆边值问题的例子 (相应于算子 $v \to -\Delta v + cv$ 的矩阵 $(a_{ij}(x))$ 对所有 $x \in \Omega$ 都等于单位矩阵), 其中相应的算子 $\mathcal{L}$ 是一致椭圆的.

最后, 变分方程 $a(u,v) = \ell(v)$ 对所有 $v \in V$, 被视为构成相应边值问题的变分形式. 这些变分方程的解 $u \in V$, 相对于经典解, 称为相应边值问题的弱解, 所谓经典解在此通常被视为在如 $\mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$ 的空间里.

习题

6.7-1 设 Ω 为 $\mathbb{R}^N$ 中的区域, 给定的函数 c, f 及 u_0 满足下述假设:

$$c \in L^\infty(\Omega), c \geqslant 0 \text{ 在 } \Omega \text{ 内几乎处处成立}, \quad f \in L^2(\Omega), u_0 \in H^1(\Omega),$$

又设 V, U 及 a 如定理 6.7-2 中给出的, 而 $\ell(v) := \int_\Omega fv dx - a(u_0, v)$ 对所有 $v \in H^1(\Omega)$.

(1) 证明相应的变分问题有唯一解 $u \in H_0^1(\Omega)$.

(2) 另外还假定 $u_0 \in H^2(\Omega)$ 及 $u \in H^2(\Omega)$, 证明 $\widetilde{u} := u + u_0$ 满足边值问题

$$-\Delta\widetilde{u} + c\widetilde{u} = f \text{ 在 } \Omega \text{ 内}, \quad \widetilde{u} = u_0 \text{ 在 } \Gamma \text{ 上}.$$

(3) 证明, 具有非齐次 Dirichlet 边界条件的同一边值问题可通过在 $H^1(\Omega)$ 的子集 $\{v \in H^1(\Omega); v = u_0 \text{ 在 } \Gamma \text{ 上}\}$ 上极小化下述泛函得到

$$J : v \in H^1(\Omega) \to J(v) = \frac{1}{2}\int_\Omega (|\nabla v|^2 + cv^2) dx - \int_\Omega fv dx.$$

6.7-2 在定理 6.7-2 中, 进一步假定对任意 $f \in L^2(\Omega), u \in H^2(\Omega)$. 证明存在常数 C 使得 $\|u\|_{2,\Omega} \leqslant C\|f\|_{0,\Omega}$ 对所有 $f \in L^2(\Omega)$.

6.7-3 设 Ω 为 $\mathbb{R}^N$ 的开子集, 而 ω_N 表示 $\mathbb{R}^N$ 的单位球的体积.

(1) 设函数 $u \in \mathcal{C}^2(\Omega)$ 使得 $-\Delta u = 0$ 在 Ω 中. 证明, 对任意 $y \in \Omega$ 及使得 $\overline{B(y;r)} \subset \Omega$ 的任意 $r > 0$, 函数 u 满足平均值性质, 即

$$u(y) = \frac{1}{N\omega_N r^{N-1}} \int_{\partial B(y;r)} u d\Gamma = \frac{1}{\omega_N r^N} \int_{B(y;r)} u dx.$$

同样地证明, 如果函数 $u \in \mathcal{C}^2(\Omega)$ 满足 $-\Delta u \geqslant 0$ 在 Ω 内, 则

$$u(y) \geqslant \frac{1}{N\omega_N r^{N-1}} \int_{\partial B(y;r)} u d\Gamma \quad \text{及} \quad u(y) \geqslant \frac{1}{\omega_N r^N} \int_{B(y;r)} u dx.$$

提示: 利用 Green 公式 $\int_{B(y;r)} \Delta u dx = \int_{\partial B(y;r)} \partial_\nu u d\Gamma$ (此式本身可立即由向量场的散度定理得到, 见 1.18 节), 计算积分时引入变量 $|x - y|$.

(2) 设函数 $u \in \mathcal{C}^2(\Omega)$ 使得 $-\Delta u \geqslant 0$ 在 Ω 内. 证明如果存在一点 $y \in \Omega$ 使得 $u(y) = \inf_{x\in\Omega} u(x)$, 则 u 是常函数; 等价地, 如果 u 不是常函数, u 不可能在 Ω 的任何一点达到其下确界.

(3) 假定 Ω 是有界的并设函数 $u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$ 满足 $-\Delta u \geqslant 0$ 在 Ω 内. 证明, u 在 $\overline{\Omega}$ 中的极小值在 $\Gamma := \partial\Omega$ 上达到, 即

$$\inf_{x\in\overline{\Omega}} u(x) = \inf_{x\in\Gamma} u(x).$$

注 (2) 及 (3) 中描述的性质分别构成上调和函数, 即在 Ω 中满足 $-\Delta u \geqslant 0$ 的函数 $u \in \mathcal{C}^2(\Omega)$, 的强及弱极小值原理. 对于一般椭圆算子类似的极小值或极大值原理, 将在 7.10 节中对更一般的情况予以讨论.

(4) 假设 Ω 是有界的, 给定函数 $f \in \mathcal{C}(\Omega)$ 与 $g \in \mathcal{C}(\Gamma)$. 证明 Dirichlet 问题

$$-\Delta u = f \text{ 在 } \Omega \text{ 内}, \quad u = g \quad \text{在 } \Gamma \text{ 上},$$

最多有一个解 $u \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega)$.

(5) 假设 Ω 是有界的, 又设 $u_\alpha \in \mathcal{C}(\overline{\Omega}) \cap \mathcal{C}^2(\Omega), \alpha \in \{1,2\}$, 是下 Dirichlet 问题的解: $-\Delta u = f$ 在 Ω 内, $u = g_\alpha, \alpha \in \{1,2\}$ 在 Γ 上. 证明

$$\sup_{x\in\overline{\Omega}} |u_1(x) - u_2(x)| \leqslant \sup_{x\in\overline{\Omega}} |g_1(x) - g_2(x)|.$$

6.7-4 设 Ω 是 $\mathbb{R}^N$ 的开子集, 函数 $u \in \mathcal{C}(\Omega)$ 使得, 对任意 $y \in \Omega$ 及任意使 $\overline{B(y;r)} \subset \Omega$ 的 $r > 0$, 有

$$u(y) = \frac{1}{N\omega_N r^{N-1}} \int_{\partial B(y;r)} u d\Gamma.$$

证明 $u \in \mathcal{C}^2(\Omega)$ 且 $-\Delta u = 0$ 在 Ω 内 (实际上, 有 $u \in \mathcal{C}^\infty(\Omega)$; 见习题 6.7-5). 这个结果与习题 6.7-3(1) 一起, 说明平均值性质完全刻画了调和函数.

提示: 以 B 表示 $\mathbb{R}^N$ 中的单位球, 利用经典的边界积分公式[32], 它对任意 $g \in \mathcal{C}(\partial B)$ 给出 Dirichlet 问题: $-\Delta u = 0$ 在 B 内, $u = g$ 在 ∂B 上, 的显式解 $w \in \mathcal{C}(B) \cap \mathcal{C}^2(B)$. 再注意, 习题 6.7-3 中问题 (2), (3), (4) 的结果同样适用于函数 $u - w$.

6.7-5 设 Ω 为 $\mathbb{R}^N$ 的开子集, w_N 表示 $\mathbb{R}^N$ 中单位球的体积. 假设 $u \in \mathcal{C}(\Omega)$ 是如习题 6.7-3(1) 中所定义的满足平均值性质的函数, 证明 $u \in \mathcal{C}^\infty(\Omega)$.

提示: 设 $(u_\varepsilon)_{\varepsilon>0}$ 是 u 的一个正则化族, 其中 $u_\varepsilon \in \mathcal{C}^\infty(\Omega_\varepsilon)$ 而 $\Omega_\varepsilon := \{x \in \Omega; \text{dist}(x, \mathbb{R}^N - \Omega) > \varepsilon\}$ (2.6 节). 证明对每个 $\varepsilon > 0$ 都有 $u = u_\varepsilon$ 在 Ω_ε 中.

6.7-6 设 Ω 为 $\mathbb{R}^N$ 的开子集, 函数 $u \in \mathcal{C}^2(\Omega)$ 满足 $-\Delta u = 0$ 在 Ω 内. 所以 $u \in \mathcal{C}^\infty(\Omega)$ (习题 6.7-4 及 6.7-5).

(1) 证明, 对任意的 $y \in \Omega$, 任意使得 $\overline{B(y;r)} \subset \Omega$ 的 $r > 0$ 以及任意整数 $k \geqslant 0$,

$$|\partial^{\boldsymbol{\alpha}} u(y)| \leqslant \frac{(2^{N+1}Nk)^k}{\omega_N r^{N+k}} \|u\|_{L^1(B(y;r))} \quad \text{对任何重指标 } \boldsymbol{\alpha}, |\boldsymbol{\alpha}| = k.$$

[32] 见, 如 Gilbarg & Trudinger [1998, 定理 2.6].

(2) 试由 (1) 推得结论: 任何满足 $-\Delta u = 0$, 在 $\mathbb{R}^N$ 内, 的有界函数 $u \in \mathcal{C}^2(\mathbb{R}^N)$ 必为常数. 这个重要性质就是著名的调和函数的 Liouville 定理[33].

(3) 证明 Liouville 定理实际上在较弱的假设下, 即调和函数 $u \in \mathcal{C}^2(\mathbb{R}^N)$ 上 (或下) 有界, 仍然成立.

提示: 利用习题 6.7-3(2).

(4) 设 Ω 为 $\mathbb{R}^N$ 的开子集, 函数 $u \in \mathcal{C}^2(\Omega)$ 在 Ω 中调和. 证明 u 在 Ω 中是解析的, 即给定任意 $y \in \Omega$, 存在 $r > 0$ 使得 $B(y;r) \subset \Omega$, 而且 u 在 $B(y;r)$ 中可展开为收敛的幂级数.

提示: 利用 (1) 中的估计, 证明, 如果 r 充分小, u 的 Taylor 级数在 $B(y;r)$ 中收敛.

6.7-7 设 Ω 是 $\mathbb{R}^N$ 中的区域.

(1) 证明存在常数 C 使得下述广义 Poincaré-Friedrichs 不等式成立:

$$\|v\|_{1,\Omega} \leqslant C\left\{\int_\Omega |\nabla v|^2 dx + \left|\int_\Omega v dx\right|^2\right\}^{1/2} \quad \text{对所有 } v \in H^1(\Omega).$$

提示: 首先证明, 若这个不等式的右端对某个 $v \in H^1(\Omega)$ 等于零, 则 $v = 0$. 再用反证法.

(2) 证明 $U := \left\{v \in H^1(\Omega); \int_\Omega v dx = 0\right\}$ 是 $H^1(\Omega)$ 的闭子空间, 而且 $|\cdot|_{1,\Omega}$ 是 U 上的一个范数, 在 U 上等价于 $\|\cdot\|_{1,\Omega}$.

(3) 设 $J(v) := \dfrac{1}{2}\int_\Omega |\nabla v|^2 dx - \ell(v), \ell(v) := \int_\Omega f v dx$ 对所有 $v \in H^1(\Omega)$. 证明 $\inf_{v\in H^1(\Omega)} J(v) > -\infty$ 意味着 $\ell(v) = 0$ 对所有常函数 v.

(4) 证明存在一个且只有一个函数 $u \in U$ 满足 $J(u) = \inf_{v\in U} J(v)$.

(5) 假设 $\ell(v) = 0$ 对所有常函数 v, 而 $u \in U \cap H^2(\Omega)$. 证明 u 满足边值问题:

$$-\Delta u = f \text{ 在 } \Omega \text{ 中}, \quad \int_\Omega u dx = 0 \text{ 及 } \partial_\nu u = 0 \text{ 在 } \Gamma \text{ 上}.$$

换言之, 在边值问题: $-\Delta u = f$ 在 Ω 中, $\partial_\nu u = 0$ 在 Γ 上, 所有可能的解 u (当 $\ell(v) = 0$ 对所有常数 v 成立时, 解存在并且在相差一个常数的意义下是确定的) 中, (3) 的极小化问题 "选取" 的 (只) 是满足 $\int_\Omega u dx = 0$ 的那一个.

提示: 首先证明, 任一函数 $\tilde{v} \in H^1(\Omega)$ 均可写为 $\tilde{v} = v + C$, 其中 $v \in U, C \in \mathbb{R}$.

6.7-8 设 Ω 是 $\mathbb{R}^N$ 中的区域.

(1) 设 $\mathcal{P}_0(\Omega)$ 表示 Ω 上所有常函数的空间. 证明半范数 $|\cdot|_{1,\Omega}$ 是商空间 $H^1(\Omega)/\mathcal{P}_0(\Omega)$ 上的范数, 等价于 $H^1(\Omega)/\mathcal{P}_0(\Omega)$ 上的商范数.

(2) 设 $\dot{w} \in H^1(\Omega)/\mathcal{P}_0(\Omega)$ 表示函数 $w \in H^1(\Omega)$ 的等价类, 令

$$\begin{aligned}
\dot{V} &:= H^1(\Omega)/\mathcal{P}_0(\Omega),\\
a(\dot{u}, \dot{v}) &:= \int_\Omega \nabla u \cdot \nabla v dx \quad \text{对所有 } \dot{u}, \dot{v} \in \dot{V} \quad \text{及}\\
\ell(\dot{v}) &:= \int_\Omega f v dx + \int_\Gamma g v d\Gamma \quad \text{对所有 } \dot{v} \in \dot{V},
\end{aligned}$$

---

[33] 冠名源自一个类似的结果, 即在单复变量的全纯 (即复解析) 函数中 "原始的" Liouville 定理, 是由 Joseph Liouville (1809—1882) 在 1847 年的一个讲座中给出的.

其中函数 $f \in L^2(\Omega)$ 与 $g \in L^2(\Gamma)$ 满足相容性条件

$$\int_\Omega f dx + \int_\Gamma g d\Gamma = 0.$$

证明, 对称双线性形式 a 是 $\dot{V}$ 强制的并且在 $\dot{V} \times \dot{V}$ 上是连续的, 线性形式 ℓ 在 $\dot{V}$ 上有明确定义且是连续的.

(3) 设 $\dot{u} \in \dot{V}$ 是变分方程 $a(\dot{u}, \dot{v}) = \ell(\dot{v})$ 对所有 $\dot{v} \in \dot{V}$ 的解 (由 (2) 知, 这个解存在且是唯一的), 假设 $u \in H^2(\Omega)$. 证明 u 满足边值问题

$$-\Delta u = f \text{ 在 } \Omega \text{ 中}, \quad \partial_\nu u = g \text{ 在 } \Gamma \text{ 上}.$$

这是算子 $-\Delta$ 的非齐次 Neumann 问题.

6.7-9 设 Ω 是 $\mathbb{R}^N$ 中的区域, Γ_1 是 $\Gamma := \partial\Omega$ 的相对开子集使得 $d\Gamma$ 测度 $\Gamma_0 > 0$, 其中 $\Gamma_0 := \Gamma - \Gamma_1$. 设函数

$$b = (b_i)_{i=1}^N \in L^\infty(\Omega; \mathbb{R}^N), \quad c \in L^\infty(\Omega) \text{ 且 } c \geqslant 0 \text{ 在 } \Omega \text{ 中几乎处处成立},$$
$$f \in L^2(\Omega), \quad g \in L^2(\Gamma_1)$$

是给定的. 令

$$V := \{v \in H^1(\Omega); v = 0 \text{ 在 } \Gamma_0 \text{ 上}\},$$
$$a(u, v) := \int_\Omega (\nabla u \cdot \nabla v + (b \cdot \nabla u)v + cuv)dx \quad \text{对所有 } u, v \in V,$$
$$\ell(v) := \int_\Omega fvdx + \int_{\Gamma_1} gvd\Gamma \quad \text{对所有 } v \in V.$$

(1) 证明, 如果 $\max_{1 \leqslant i \leqslant N} \|b_i\|_{L^\infty(\Omega)}$ 充分小, 则变分方程 $a(u, v) = \ell(v)$ 对所有 $v \in V$, 存在唯一解 $u \in V$.

(2) 如果 $u \in H^2(\Omega)$, u 满足何种边值问题?

6.7-10 本题的目标是分析一个模型奇摄动问题[34] 的解当 $\varepsilon \to 0$ 时的性态, 所谓奇摄动是指当 $\varepsilon > 0$ 时被参数化的问题, 当 $\varepsilon \to 0$ 时变成 (在某种意下) "奇异的".

(1) 设 Ω 是 $\mathbb{R}^N$ 的有界开子集, $f \in L^2(\Omega)$ 是给定的, 设 $u_\varepsilon \in H_0^1(\Omega)$ 对每个 $\varepsilon > 0$ 表示以下问题的解:

$$-\varepsilon \Delta u_\varepsilon + u_\varepsilon = f \text{ 在 } \Omega \text{ 中}, \quad u_\varepsilon = 0 \text{ 在 } \Gamma \text{ 上}.$$

证明当 $\varepsilon \to 0$ 时 $u_\varepsilon \to f$ 在 $L^2(\Omega)$ 中.

提示: 证明函数族 $(\sqrt{\varepsilon} u_\varepsilon)_{\varepsilon > 0}$ 在 $H_0^1(\Omega)$ 中有界, 函数族 $(u_\varepsilon)_{\varepsilon > 0}$ 在 $L^2(\Omega)$ 中有界.

(2) 在 $f \in H_0^1(\Omega)$ 的附加假设下, 证明当 $\varepsilon \to 0$ 时 $u_\varepsilon \to f$ 在 $H_0^1(\Omega)$ 中.

6.8 四阶线性边值问题的实例; 重调和与板问题

整个这一节中, $\mathbb{R}^N$ 中区域 Ω 的边界均以 Γ 表示.

[34] 关于奇摄动问题的参考文献在文献注释中给出.

在前面一节中, 空间 V 是 Sobolev 空间 $H^1(\Omega)$ 的子空间, 而在我们现在考虑的例子中, 空间 V 将视为 $H^2(\Omega)$ 的子空间. 作为对于本节中我们第一个实例的准备, 先证明两个简单的预备性结果. 首先回忆一下, 如果开集 $\Omega \subset \mathbb{R}^N$ 具有有限宽度, 半范数 $|\cdot|_{2,\Omega} : H^2(\Omega) \to \mathbb{R}$ 就成为空间 $H_0^2(\Omega)$ 上等价于 $\|\cdot\|_{2,\Omega}$ 的范数 (定理 6.5-2(b)). 下一个结果说明, 在这种情况下, 空间 $H_0^2(\Omega)$ 可被赋予另一种等价的范数.

定理 6.8-1 (a) 设 Ω 为 $\mathbb{R}^N$ 的开子集. 则

$$|v|_{2,\Omega} = \|\Delta v\|_{0,\Omega} \quad 对所有\ v \in H_0^2(\Omega).$$

(b) 如果 Ω 具有有限宽度, 则半范数 $v \to \|\Delta v\|_{0,\Omega}$ 成为空间 $H_0^2(\Omega)$ 上等价于范数 $\|\cdot\|_{2,\Omega}$ 的范数.

证明 只需对 $H_0^2(\Omega)$ 的稠密子空间 $\mathcal{D}(\Omega)$ 中的函数证明等式 $|v|_{2,\Omega} = \|\Delta v\|_{0,\Omega}$. 注意到, 由定义有

$$|v|_{2,\Omega}^2 = \int_\Omega \left(\sum_{i=1}^N |\partial_{ii} v|^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^N |\partial_{ij} v|^2 \right) dx,$$

$$\|\Delta v\|_{0,\Omega}^2 = \int_\Omega \left(\sum_{i=1}^N |\partial_{ii} v|^2 + \sum_{\substack{i,j=1 \\ i \neq j}}^N \partial_{ii} v \partial_{jj} v \right) dx,$$

而由定理 6.3-1 有

$$\int_\Omega |\partial_{ij} v|^2 dx = -\int_\Omega \partial_i v \partial_{ijj} v dx = \int_\Omega \partial_{ii} v \partial_{jj} v dx \quad 对所有\ v \in \mathcal{D}(\Omega),$$

立即就得到该等式的正确性. 所以 (a) 得证. 结合 (a) 与当 $m = 2$ 时的定理 6.5-2(b), (b) 立即得证. □

第二个预备性的结果构成 Sobolev 空间中的另一个 Green 公式. 出现在该公式中的算子 $\Delta^2 := \sum_{i,j=1}^N \partial_{iijj}$ (其作用于定义在 Ω 内的函数上) 称为重调和算子.

定理 6.8-2 设 Ω 是 $\mathbb{R}^N$ 中的区域, $\nu = (\nu_i)_{i=1}^N$ 表示沿边界 Γ 的单位外法向量场. 则下述 Green 公式成立:

$$\int_\Omega \Delta u \Delta v dx = \int_\Omega (\Delta^2 u) v dx - \int_\Gamma (\partial_\nu \Delta_u) v d\Gamma$$
$$+ \int_\Gamma \Delta u \partial_\nu v d\Gamma \quad 对所有\ u \in H^4(\Omega), v \in H^2(\Omega),$$

其中

$$\Delta^2 u := \Delta(\Delta u) = \sum_{i,j=1}^N \partial_{iijj} u \in L^2(\Omega),$$

而 $\partial_\nu v \in L^2(\Gamma)$ 表示 $v \in H^2(\Omega)$ 的外法向导数 (定理 6.7-1).

证明 由定理 6.7-1(a),

$$\begin{aligned}\int_\Omega \sum_{i=1}^N \partial_i v \partial_i w dx &= -\int_\Omega (\Delta v) w dx + \int_\Gamma (\partial_\nu v) w d\Gamma \\ &= -\int_\Omega v \Delta w dx + \int_\Gamma v \partial_\nu w d\Gamma \quad \text{对所有 } v, w \in H^2(\Omega).\end{aligned}$$

所以

$$\int_\Omega (v\Delta w - (\Delta v) w) dx = \int_\Gamma (v \partial_\nu w - (\partial_\nu v) w) dx \quad \text{对所有 } v, w \in H^2(\Omega).$$

只需在这个 Green 公式中以 Δu 取代 w, 就得要证的公式. □

我们现在考虑一个在空间 $H_0^2(\Omega)$ 中的变分问题.

定理 6.8-3 设 Ω 为 $\mathbb{R}^N$ 中具有有限宽度的开子集, 函数 $f \in L^2(\Omega)$ 是给定的, 又令

$$\begin{aligned} V = U :&= H_0^2(\Omega), \\ a(u, v) :&= \int_\Omega \Delta u \Delta v dx \quad \text{对所有 } u, v \in V, \\ \ell(v) :&= \int_\Omega f v dx \quad \text{对所有 } v \in V. \end{aligned}$$

则存在唯一的函数 $u \in H_0^2(\Omega)$, 它极小化由下式定义的泛函 $J : H_0^2(\Omega) \to \mathbb{R}$:

$$\begin{aligned} J(v) :&= \frac{1}{2} a(v, v) - \ell(v) \\ &= \frac{1}{2} \int_\Omega |\Delta v|^2 dx - \int_\Omega f v dx \quad \text{对所有 } v \in H_0^2(\Omega), \end{aligned}$$

或等价地满足变分方程

$$\int_\Omega \Delta u \Delta v dx = \int_\Omega f v dx \quad \text{对所有 } v \in H_0^2(\Omega).$$

u 满足下述边值问题:

$$\Delta^2 u = f \text{ 在 } \Omega \text{ 内}, \quad u = \partial_\nu u = 0 \text{ 在 } \Gamma \text{ 上},$$

其中 Ω 中的偏微分方程理解为空间 $\mathcal{D}'(\Omega)$ 中的等式, 在 Ω 是区域的附加假设下, 边界条件理解为空间 $L^2(\Gamma)$ 中的等式.

证明 对称双线性形式 $a : H^2(\Omega) \times H^2(\Omega) \to \mathbb{R}$ 是连续的, 此因

$$\begin{aligned} |a(u, v)| &\leqslant \|\Delta u\|_{0,\Omega} \|\Delta v\|_{0,\Omega} \\ &\leqslant N \|u\|_{2,\Omega} \|v\|_{2,\Omega} \quad \text{对所有 } u, v \in H^2(\Omega), \end{aligned}$$

又是 $H_0^2(\Omega)$ 强制的, 此因

$$a(v,v)=\|\Delta v\|_{0,\Omega}^2 \quad 对所有\ v\in H_0^2(\Omega),$$

而且 $v\in H^2(\Omega)\to\|\Delta v\|_{0,\Omega}$ 在 $H_0^2(\Omega)$ 上是等价于 $\|\cdot\|_{2,\Omega}$ 的范数 (定理 6.8-1).

因此定理 6.1-1 中的所有假设都满足 (线性形式 $\ell: H_0^2(\Omega)\to\mathbb{R}$ 显然是连续的), 故得存在一个且只有一个函数极小化所示泛函 J 或等价地满足所示变分方程 (定理 6.1-2).

因为 $\int_\Omega \Delta u\Delta v dx=\langle\Delta^2 u, v\rangle$ 对所有 $v\in\mathcal{D}(\Omega)$, 其中 $\Delta^2 u$ 在此视为一个分布, 方程 $a(u,v)=\ell(v)$ 对所有 $v\in V$, 意味着

$$\langle\Delta^2 u, v\rangle=\int_\Omega fvdx \quad 对所有\ v\in\mathcal{D}(\Omega)$$

(此因 $\mathcal{D}(\Omega)\subset H_0^2(\Omega)$), 所以

$$\Delta^2 u=f \quad 在\ \mathcal{D}'(\Omega)\ 中.$$

当 Ω 是区域时, 空间 $H_0^2(\Omega)$ 可表示为

$$H_0^2(\Omega)=\{v\in H^2(\Omega); v=\partial_\nu v=0\ 在\ \Gamma\ 上\}$$

(定理 6.6-5(d)), 这说明函数 $u\in H_0^2(\Omega)$ 满足边界条件 $u=0$ 和 $\partial_\nu u=0$ 在 Γ 上, 它们被视为空间 $L^2(\Gamma)$ 中的等式. □

边值问题

$$\Delta^2 u=f\ 在\ \Omega\ 内, \quad u=\partial_\nu u=0\ 在\ \Gamma\ 上$$

称为重调和问题.

下一个变分问题形成于 $\mathbb{R}^2$ 中的一区域上, 相应地我们也需要一些在二维情况下建立起来的预备性结果. 第一个预备性结果 (其证明类似于定理 6.7-5 中的相关内容, 作为习题留给读者. 见习题 6.8-1) 将被用于确立相关双线性形式的椭圆性.

定理 6.8-4 设 Ω 是 $\mathbb{R}^2$ 中的区域, Γ_0 是边界 Γ 的 $d\Gamma$ 可测子集且满足

$$d\Gamma\ 测度\ \Gamma_0>0,$$

又令

$$V:=\{v\in H^2(\Omega);\ v=\partial_\nu v=0\ 在\ \Gamma_0\ 上\}.$$

则空间 V 是 $H^2(\Omega)$ 的闭子空间, 且存在常数 C 使得

$$|v|_{2,\Omega}\leqslant\|v\|_{2,\Omega}\leqslant C|v|_{2,\Omega} \quad 对所有\ v\in V.$$

下面两个预备性结果在验证出现在关联边值问题中的边界条件时起到实质性作用. 其中第一个可说是定理 6.7-3 的 “$H^2(\Omega)$ 版”.

定理 6.8-5 设 Ω 是 $\mathbb{R}^2$ 中的区域, Γ_1 是 Γ 的 $\mathcal{C}^{1,1}$ 类相对开子集, 又设 $w_0, w_1 \in L^2(\Gamma_1)$ 是两个函数, 它们满足

$$\int_{\Gamma_1} w_0 v d\Gamma + \int_{\Gamma_1} w_1 \partial_\nu v d\Gamma = 0$$

$$\text{对所有 } v \in V := \{v \in H^2(\Omega); v = \partial_\nu v = 0 \text{ 在 } \Gamma - \Gamma_1 \text{ 上}\}.$$

则 $w_0 = w_1 = 0$.

下一个结果构成 Sobolev 空间中的另一个 Green 公式. 在陈述该公式之前, 我们需要几个 $\mathbb{R}^2$ 中区域特有的定义及符号.

设 Ω 表示 $\mathbb{R}^2$ 中的区域, $\nu = (\nu_\alpha)_{\alpha=1}^2$ 表示沿 Γ 的单位外法向量场. 沿 Γ 的单位切向量场 $\tau = (\tau_\alpha)_{\alpha=1}^2$ 由下式定义

$$\tau_1 = -\nu_2, \quad \tau_2 = \nu_1.$$

所以与 ν 一样, 向量场 τ 也是沿 Γ 上 $d\Gamma$ 几乎处处确定的.

除了法向微分算子 ∂_ν 之外, 我们由以下诸式定义其他沿 Γ 的边界微分算子 $\partial_\tau, \partial_{\nu\tau}, \partial_{\tau\tau}$:

$$\partial_\tau v := \sum_{\alpha=1}^{2} \tau_\alpha \partial_\alpha v, \quad \partial_{\nu\tau} v := \sum_{\alpha,\beta=1}^{2} \nu_\alpha \tau_\beta \partial_{\alpha\beta} v,$$

$$\partial_{\tau\tau} v := \sum_{\alpha,\beta=1}^{2} \tau_\alpha \tau_\beta \partial_{\alpha\beta} v$$

对于充分光滑的函数 v. 顺便提一下, 在单位切向量确定的那些边界点上, 将函数 v 在 Γ 上的限制视为沿 Γ 的曲线横坐标的函数, $\partial_\tau v$ 与其一阶导数一致, 但一般来说, $\partial_{\tau\tau} v$ 与这个限制的二阶导数并不相同.

为简洁起见, 在本节的其余部分, 希腊字母指标的取值范围默认为在集合 $\{1, 2\}$ 中, 而且对希腊字母指标采用求和简约, 如以 $m_{\alpha\beta}\partial_{\alpha\beta} v$ 表示 $\sum_{\alpha,\beta=1}^2 m_{\alpha\beta}\partial_{\alpha\beta} v$ 等.

定理 6.8-6 设 Ω 为 $\mathbb{R}^2$ 中的区域. 则下述 Green 公式成立:

$$\int_\Omega m_{\alpha\beta}\partial_{\alpha\beta} v dx = \int_\Omega (\partial_{\alpha\beta} m_{\alpha\beta}) v dx - \int_\Gamma ((\partial_\alpha m_{\alpha\beta})\nu_\beta + \partial_\tau(m_{\alpha\beta}\nu_\alpha\tau_\beta)) v d\Gamma$$

$$+ \int_\Gamma m_{\alpha\beta}\nu_\alpha\nu_\beta \partial_\nu v d\Gamma \quad \text{对所有 } m_{\alpha\beta} \in H^2(\Omega), v \in H^2(\Omega).$$

证明　接连两次应用 Sobolev 空间中的基本 Green 公式 (定理 6.6-7) 给出

$$\int_\Omega m_{\alpha\beta}\partial_{\alpha\beta} v dx = -\int_\Omega (\partial_\alpha m_{\alpha\beta})\partial_\beta v dx + \int_\Gamma m_{\alpha\beta}\nu_\alpha \partial_\beta v d\Gamma$$

$$= \int_\Omega (\partial_{\alpha\beta} m_{\alpha\beta}) v dx - \int_\Gamma (\partial_\alpha m_{\alpha\beta})\nu_\beta v d\Gamma + \int_\Gamma m_{\alpha\beta}\nu_\alpha \partial_\beta v d\Gamma.$$

边界算子 ∂_ν 与 ∂_τ 的定义意味着沿 Γ, v 的每个偏导数都可以写成 $\partial_\beta v = \nu_\beta \partial_\nu v + \tau_\beta \partial_\tau v$. 所以最后一个 Γ 上的积分可写为

$$\int_\Gamma m_{\alpha\beta}\nu_\alpha \partial_\beta v d\Gamma = \int_\Gamma m_{\alpha\beta}\nu_\alpha \nu_\beta \partial_\nu v d\Gamma + \int_\Gamma m_{\alpha\beta}\nu_\alpha \tau_\beta \partial_\tau v d\Gamma,$$

再注意到 (习题 6.8-2)

$$\int_\Gamma m_{\alpha\beta}\nu_\alpha \tau_\beta \partial_\tau v d\Gamma = -\int_\Gamma (\partial_\tau (m_{\alpha\beta}\nu_\alpha \tau_\beta)) v d\Gamma,$$

就得到要证的 Green 公式. □

我们现在考虑第二个实例. 回忆一下, 符号 $\delta_{\alpha\beta}$ 表示 Kronecker 记号.

定理 6.8-7 设 Ω 为 $\mathbb{R}^2$ 中的区域, Γ_1 是 Γ 的 $\mathcal{C}^{1,1}$ 类相对开子集, 且

$$d\Gamma \text{ 测度 } \Gamma_0 > 0, \quad \text{其中 } \Gamma_0 := \Gamma - \Gamma_1,$$

又设

$$0 < \nu < 1, \quad f \in L^2(\Omega),$$

是给定的常数及给定函数, 最后设

$$\begin{aligned} V = U :&= \{v \in H^2(\Omega); v = \partial_\nu v = 0 \text{ 在 } \Gamma_0 \text{ 上}\}, \\ a(u,v) :&= \int_\Omega (\nu \Delta u \Delta v + (1-\nu)\partial_{\alpha\beta} u \partial_{\alpha\beta} v) dx \quad \text{对所有 } u, v \in V, \\ \ell(v) :&= \int_\Omega f v dx \quad \text{对所有 } v \in V. \end{aligned}$$

则存在唯一的函数 $u \in V$, 它在空间 V 上极小化由下式定义的泛函 $J: V \to \mathbb{R}$:

$$\begin{aligned} J(v) :&= \frac{1}{2} a(v,v) - \ell(v) \\ &= \frac{1}{2}\int_\Omega (\nu |\Delta v|^2 + (1-\nu)\partial_{\alpha\beta} v \partial_{\alpha\beta} v) dx - \int_\Omega f v dx \end{aligned}$$

对所有 $v \in V$ 或等价地满足

$$\int_\Omega (\nu \Delta u \Delta v + (1-\nu)\partial_{\alpha\beta} u \partial_{\alpha\beta} v) dx = \int_\Omega f v dx \quad \text{对所有 } v \in V.$$

再进一步假定 $u \in H^4(\Omega)$, 而函数 $m_{\alpha\beta}(u) \in H^2(\Omega)$ 由下式定义:

$$m_{\alpha\beta}(u) := \nu \Delta u \delta_{\alpha\beta} + (1-\nu)\partial_{\alpha\beta} u.$$

则 u 满足边值问题:

$$\begin{aligned} \partial_{\alpha\beta} m_{\alpha\beta}(u) &= f \quad \text{在 } \Omega \text{ 内}, \\ u = \partial_\nu u &= 0 \quad \text{在 } \Gamma_0 \text{ 上}, \\ m_{\alpha\beta}(u)\nu_\alpha \nu_\beta = (\partial_\alpha m_{\alpha\beta}(u))\nu_\beta + \partial_\tau (m_{\alpha\beta}(u)\nu_\alpha \tau_\beta) &= 0 \quad \text{在 } \Gamma_1 \text{ 上}. \end{aligned}$$

证明 对称的双线性形式 $a: V \times V \to \mathbb{R}$ 及线性形式 $\ell: V \to \mathbb{R}$ 显然是连续的. 根据定理 6.8-4, 双线性形式还是 V 强制的, 这是因为 $0 < \nu < 1$ 且

$$a(v,v) \geqslant (1-\nu)|v|_{2,\Omega}^2 \quad 对所有\ v \in V.$$

所以存在唯一的函数 u, 它在空间 V 上极小化所示泛函 J 或等价地满足所示的变分方程.

为了验证相应的边值问题, 首先注意, 变分方程的左端也可以写成

$$\int_\Omega (\nu \Delta u \Delta v + (1-\nu)\partial_{\alpha\beta} u \partial_{\alpha\beta} v) dx = \int_\Omega m_{\alpha\beta}(u) \partial_{\alpha\beta} v dx,$$

其中 $m_{\alpha\beta}(u) := \nu \Delta u \delta_{\alpha\beta} + (1-\nu)\partial_{\alpha\beta} u$. 假定 $u \in H^4(\Omega)$, 这样就有 $m_{\alpha\beta}(u) \in H^2(\Omega)$; 利用定理 6.8-6 中 Green 公式并考虑到 $v = \partial_\nu v = 0$ 在 Γ_0 上, 变分方程 $a(u,v) = \ell(v)$ 对所有 $v \in V$, 就变成

$$\begin{aligned}\int_\Omega (\partial_{\alpha\beta} m_{\alpha\beta}(u) - f) v dx = &\int_{\Gamma_1} \{(\partial_\alpha m_{\alpha\beta}(u))\nu_\beta + \partial_\tau (m_{\alpha\beta}(u)\nu_\alpha \tau_\beta)\} v d\Gamma \\ &- \int_{\Gamma_1} m_{\alpha\beta}(u)\nu_\alpha \nu_\beta \partial_\nu v d\Gamma \quad 对所有\ v \in V.\end{aligned}$$

特别地有

$$\int_\Omega (\partial_{\alpha\beta} m_{\alpha\beta}(u) - f) v dx = 0 \quad 对所有\ v \in \mathcal{D}(\Omega),$$

这意味着在 $L^2(\Omega)$ 中 $\partial_{\alpha\beta} m_{\alpha\beta}(u) = f$. 注意到已得的这个方程, 剩下来的就是

$$\begin{aligned}&\int_{\Gamma_1} \{(\partial_\alpha m_{\alpha\beta}(u))\nu_\beta + \partial_\tau (m_{\alpha\beta}(u)\nu_\alpha \tau_\beta)\} v d\Gamma \\ &- \int_{\Gamma_1} m_{\alpha\beta}(u)\nu_\alpha \nu_\beta \partial_\nu v d\Gamma = 0 \quad 对所有\ v \in V,\end{aligned}$$

由定理 6.8-5 知, 这意味着所示在 Γ_1 上的边界条件作为 $L^2(\Gamma_1)$ 中的等式满足. 最后, $u \in V$ 意味着 $u = \partial_\nu u = 0$ 在 Γ_0 上. □

出现在定理 6.8-7 中的数据 $V, a(\cdot,\cdot)$ 及 ℓ 相应于线性弹性板的 Kirchhoff-Love 理论中弯曲方程的变分形式: 未知函数 u 表示常厚度 e 的线性弹性板在单位面积密度 $F = \frac{1}{12} E e^3 f/(1-\nu^2)$ 的垂直力作用下的垂直位移. 常数 $E = \mu(3\lambda + 2\mu)/(\lambda+\mu)$ 与 $\nu = \frac{1}{2}\lambda/(\lambda+\mu)$ 分别为构成板的弹性材料的 Young 氏模量与 Poisson 系数, $\lambda \geqslant 0$ 与 $\mu > 0$ 是同一材料的 Lamé 常数; 所以 Poisson 系数满足 $0 < \nu < 1/2$. 当 $f = 0$ 时, 板位于坐标为 (x_1, x_2) 的 "水平" 平面内 (图 6.8-1). 包含在空间 V 定义中的边界条件 $u = \partial_\nu u = 0$ 在 Γ_0 上, 反映了板在 Γ_0 处是紧固的.

板的未知垂直位移极小化板的能量 $J: V \to \mathbb{R}$, 其由下式定义:

$$J(v) := \frac{1}{2} \int_\Omega (\nu |\Delta v|^2 + (1-\nu)\partial_{\alpha\beta} v \partial_{\alpha\beta} v) dx - \int_\Gamma f v dx \quad 对所有\ v \in V.$$

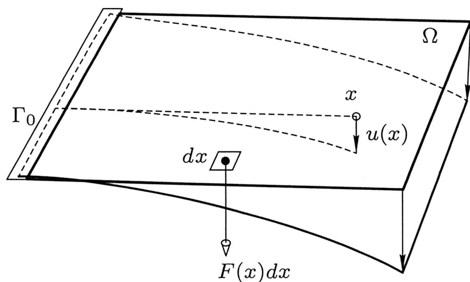


图 6.8-1 板问题: 未知函数 $u:\overline{\Omega}\subset\mathbb{R}^2\to\mathbb{R}$ 表示线性弹性板在单位面积密度 F 的垂直力作用下的垂直位移. 板在无外力的情况占据集合 $\overline{\Omega}$, 其边界 Γ 的一部分 Γ_0 是紧固的

注意, 由定理 6.8-1(a), 全部边界都紧固的板 (在这种情况, $V=H_0^2(\Omega)$) 的能量有较简单的形式

$$J(v):=\frac{1}{2}\int_\Omega|\Delta v|^2dx-\int_\Omega fvdx\quad 对所有\ v\in H_0^2(\Omega),$$

即与相应于重调和问题 (定理 6.8-3) 的泛函相同.

同样的重调和问题也是流体力学中一类特定问题的数学模型: 可以证明, 在单连通区域 $\Omega\subset\mathbb{R}^2$ 中, 不可压缩黏性流体的 Stokes 方程 (6.14 节) 的解可化为求解上述重调和问题, 其中未知函数 u 是适当的流函数.

注 固定板问题的变分形式中的表达式可以严格地通过应用渐近分析方法 (当板的厚度趋向零时) 于三维线性弹性边值问题的变分形式[35] (将在 6.16 节中讨论) 予以验证.

如果 $u\in H^4(\Omega)$, 方程 $\partial_{\alpha\beta}m_{\alpha\beta}(u)=f$ 在 Ω 内, 也可以写为 $\Delta^2u=f$ 在 Ω 内 (因为 $\partial_{\alpha\beta}m_{\alpha\beta}(u)=\Delta^2u$). 所以出现在定理 6.8-3 与 6.8-7 中的变分问题为我们提供两个非常有趣的变分问题的实例, 它们具有不同的双线性形式, 然而却导致 Ω 中同样的偏微分方程 (这方面, 也可见习题 6.8-3). 只是当 $d\Gamma$ 测度 $\Gamma_1>0$ 时, 边界条件不同.

习题

6.8-1 设 Ω 为 $\mathbb{R}^2$ 中的区域, $\Gamma_0\subset\partial\Omega$ 且 $d\Gamma$ 测度 $\Gamma_0>0$.

(1) 证明空间 $V:=\{v\in H^2(\Omega);v=\partial_\nu v=0\ 在\ \Gamma_0\ 上\}$ 在 $H^2(\Omega)$ 中是闭的且 $|\cdot|_{2,\Omega}$ 是 V 上的范数.

提示: 由定理 6.3-4 推得, 如果函数 $v\in H^2(\Omega)$ 满足 $|v|_{2,\Omega}=0$, 则存在常数 $a_i, 0\leqslant i\leqslant 2$, 使得 $v(x)=a_0+a_1x_1+a_2x_2$ 对所有 $x=(x_i)_{i=1}^2\in\Omega$; 然后再证明, 如果进一步还有 $\nu=\partial_\nu v=0$ 在 Γ_0 上且 $d\Gamma$ 测度 $\Gamma_0>0$, 则 $v=0$.

[35] 线性板模型的这类验证在 Ciarlet [1997, 第 1 章] 中有详尽的讨论.

(2) 假定存在函数 $v_k \in V$ 的序列 $(v_k)_{k=1}^{\infty}$ 满足

$$\|v_k\|_{2,\Omega} = 1 \quad \text{对所有 } k \quad \text{且} \quad \lim_{k\to\infty} |v_k|_{2,\Omega} = 0.$$

结合 Rellich-Kondrachov 定理 (定理 6.6-3) 与问题 (1), 证明这个假设导致矛盾. 所以存在常数 C 使得 $\|v\|_{2,\Omega} \leqslant C|v|_{2,\Omega}$ 对所有 $v \in V$.

6.8-2 设 Ω 是 $\mathbb{R}^2$ 中的区域, $\Gamma := \partial\Omega$. 证明

$$\int_\Gamma m_{\alpha\beta}\nu_\alpha\tau_\beta\partial_\tau v d\Gamma = -\int_\Gamma (\partial_\tau(m_{\alpha\beta}\nu_\alpha\tau_\beta))v d\Gamma \quad \text{对所有函数 } m_{\alpha\beta} \in H^2(\Omega), v \in H^2(\Omega).$$

6.8-3 设 Ω 是 $\mathbb{R}^2$ 中的区域, $\Gamma := \partial\Omega$.

(1) 证明下述 Sobolev 空间中的 Green 公式成立:

$$\begin{aligned}&\int_\Omega (2\partial_{12}u\partial_{12}v - \partial_{11}u\partial_{22}v - \partial_{22}u\partial_{11}v)dx\\ &= \int_\Gamma (-\partial_{\tau\tau}u\partial_\nu v + \partial_{\nu\tau}u\partial_\tau v)d\Gamma\end{aligned}$$

对所有 $u \in H^3(\Omega), v \in H^2(\Omega)$.

(2) 对任意的 $\nu \in \mathbb{R}$, 证明

$$\begin{aligned}&\int_\Omega (\nu\Delta u\Delta v + (1-\nu)\partial_{\alpha\beta}u\partial_{\alpha\beta}v)dx\\ &= \int_\Omega (\Delta u\Delta v + (1-\nu)(2\partial_{12}u\partial_{12}v - \partial_{11}u\partial_{22}v - \partial_{22}u\partial_{11}v))dx\end{aligned}$$

对所有 $u, v \in H^2(\Omega)$. 综合这一结果与问题 (1) 中的 Green 公式, 证明, 如果定理 6.8-7 中变分问题的解 u 属于空间 $H^4(\Omega)$, 则 u 满足偏微分方程 $\Delta^2 u = f$ 在 Ω 内.

6.8-4 设 ω 是 $\mathbb{R}^2$ 中的区域. 我们在定理 6.8-1 中已证明, $\eta \to \|\Delta\eta\|_{0,\omega}$ 是空间 $H_0^2(\omega)$ 上等价于 $\|\cdot\|_{2,\omega}$ 的范数.

(1) 假定 ω 具有光滑边界 γ, 又设 $\gamma_0 \subset \gamma$ 且 $0 <$ γ_0 的长度 $<$ γ 的长度. 证明 $\eta \to \|\Delta\eta\|_{0,\omega}$ 还是空间 $V(\omega) := \{\eta \in H^2(\omega); \eta = \partial_\nu\eta = 0 \text{ 在 } \gamma_0 \text{ 上}\}$ 上的范数.

(2) 这个范数在 $V(\omega)$ 上是否等价于范数 $\|\cdot\|_{2,\omega}$?

6.8-5 设 Ω 为 $\mathbb{R}^N$ 中的区域, $\Gamma := \partial\Omega$, 又设

$$V := \{v \in H^2(\Omega); v = 0 \text{ 在 } \Gamma \text{ 上}\}.$$

(1) 证明 $v \to \|\Delta u\|_{0,\Omega}$ 是 V 上的范数.

(2) 这个范数在 V 上是否等价于范数 $\|\cdot\|_{2,\Omega}$?

6.9 与变分不等式相应的非线性边值问题的实例; 障碍问题

在这一节中, 我们要讨论的变分问题是以变分不等式的形式出现的, 这个问题起源于在一集合上极小化一个二次泛函, 但该集合不是向量空间的情况 (定理 6.1-2). 我

们从一个特定的实例开始, 这个例子是薄膜问题 (6.7 节) 的有趣的变形. 回忆一下, 由下式定义的泛函 $J : H_0^1(\Omega) \to \mathbb{R}$,

$$J(v) := \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \int_\Omega fv dx \quad 对所有\ v \in H_0^1(\Omega),$$

表示弹性薄膜的能量, 薄膜通过水平面 $\mathbb{R}^2$ 上一区域 Ω 的边界并且受到密度 $F = \tau f$ 的垂直力的作用, 其中 $f \in L^2(\Omega)$, 而 τ 表示薄膜的张力 (6.7 节).

薄膜的障碍问题仍然在于求其平衡位置, 只是要在附加的假设下进行, 即如图 6.9-1 所示, 它必须位于由函数 $\chi : \overline{\Omega} \to \mathbb{R}$ 表示的 "障碍" 之上 (自然地, 假定函数 $\chi \leqslant 0$ 在 Γ 上). 因此, 现在期望未知的垂直位移 u 不是在整个空间 $H_0^1(\Omega)$, 而是在集合 $U := \{v \in H_0^1(\Omega); v \geqslant \chi$ 几乎处处在 Ω 中$\}$ 上极小化同样的泛函 J.

我们现在来确立这种极小函数 $u \in U$ 的存在性及唯一性, 而且也如通常所作, 在附加的正则性假设下 (然而, 要特别注意假设的合理性, 见证明后的简短讨论), 导出 u 所满足的非线性边值问题.

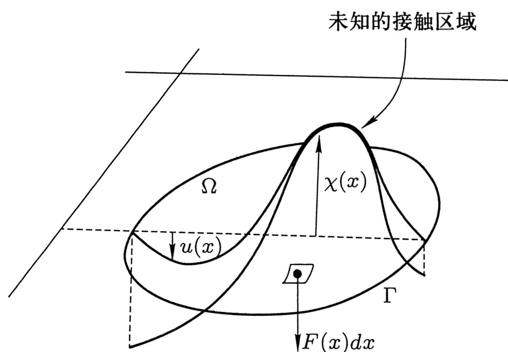


图 6.9-1 障碍问题: 薄膜必须位于由一个函数 $\chi : \overline{\Omega} \to \mathbb{R}$ 表示的 "障碍" 之上. 该图最早出现在: P. G. Ciarlet [1978]: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam

定理 6.9-1 设 Ω 是 $\mathbb{R}^2$ 中的区域, 函数

$$\chi \in H^1(\Omega) \cap \mathcal{C}(\overline{\Omega})\ 且\ \chi|_\Gamma \leqslant 0, \quad f \in L^2(\Omega)$$

是给定的, 又设

$$V := H_0^1(\Omega)\ 而\ U := \{v \in H_0^1(\Omega); v \geqslant \chi\ 几乎处处在\ \Omega\ 内\},$$

$$a(u,v) := \int_\Omega \nabla u \cdot \nabla v dx \quad 对所有\ u, v \in V,$$

$$\ell(v) := \int_\Omega fv dx \quad 对所有\ v \in V.$$

则存在唯一的函数 $v \in U$ 在集合 U 上极小化由下式定义的泛函 $J : V \to \mathbb{R}$,

$$J(v) := \frac{1}{2}a(v,v) - \ell(v) = \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \int_\Omega fv dx$$

或等价地满足变分不等式

$$\int_\Omega \nabla u \cdot \nabla(v-u)dx \geqslant \int_\Omega f(v-u)dx \quad \text{对所有 } v \in U.$$

此外, 以这种方式定义的映射

$$f \in L^2(\Omega) \to u \in U \subset H_0^1(\Omega)$$

是非线性并且 Lipschitz 连续的.

进一步假定 $u \in H^2(\Omega)$. 则 u 满足下述边值问题:

$$\begin{aligned} -\Delta u = f &\quad \text{在 } \Omega^+ := \{y \in \Omega; u(y) > \chi(y)\} \text{ 内}, \\ -\Delta u \geqslant f &\quad \text{在 } \Omega^0 := \{y \in \Omega; u(y) = \chi(y)\} := \Omega - \Omega^+ \text{ 内}, \\ u \geqslant \chi &\quad \text{在 } \overline{\Omega} \text{ 内}, \\ u = 0 &\quad \text{在 } \Gamma \text{ 上}. \end{aligned}$$

证明　(i) 除了关于集合 U 的假设外, 定理 6.1-1 中的所有假设都满足 (见定理 6.7-2 的证明).

给定函数 $v \in H^1(\Omega)$, 函数 $\max\{0, v\}$ 也属于空间 $H^1(\Omega)$[36]. 因此可得, 如果另外还有 $\operatorname{tr} v|_\Gamma \leqslant 0\ d\Gamma$ 几乎处处在 Γ 上成立 (作为 $L^2(\Gamma)$ 中的函数), 函数 $\max\{0, v\}$ 属于空间 $H_0^1(\Omega)$. 所以 $H_0^1(\Omega)$ 的子集 U 是非空的, 因为它包含函数 $\max\{0, \chi\}$. 它也是凸的, 因为

$$\lambda v + (1-\lambda)w \geqslant \lambda\chi + (1-\lambda)\chi = \chi \quad \text{几乎处处在 } \Omega \text{ 内},$$
$$\text{对所有 } v, w \in U \text{ 及所有 } 0 < \lambda < 1,$$

且又是闭的: 设函数 $v_k \in U, k \geqslant 1$, 而 $v \in H_0^1(\Omega)$ 使得 $\|v_k - v\|_{1,\Omega} \to 0$ 当 $k \to \infty$ 时, 因此更有 $\|v_k - v\|_{0,\Omega} \to 0$ 当 $k \to \infty$ 时. 所以有一个子列 $(v_{\sigma(k)})_{k=1}^\infty$ 在 Ω 内几乎处处点点收敛于 v (定理 3.4-3). 这就有

$$v(x) = \lim_{k\to\infty} v_{\sigma(k)}(x) \geqslant \chi(x) \quad \text{对几乎所有} x \in \Omega.$$

所以存在唯一的函数 $u \in U$, 它在集合 U 上极小化所示泛函 J (定理 6.1-1) 或等价地满足所示的变分不等式 (定理 6.1-2).

线性映射 $f \in L^2(\Omega) \to \ell \in V'$ 是连续的, 此因

$$\|\ell\|_{V'} = \sup_{\substack{v \in V \\ v \neq 0}} \frac{|\ell(v)|}{\|v\|_{1,\Omega}} \leqslant \|f\|_{0,\Omega} \quad \text{对所有 } f \in L^2(\Omega),$$

[36] 这个结果 (它并不是平凡的) 的一个证明, 可见如:

G. Stampacchia [1965]: Equations Elliptiques du Second Ordre à Coefficients Discontinus. Presses de l'Université de Montréal, Montréal, Que.

而且非线性映射 $\ell \in V' \to u \in U \subset H_0^1(\Omega)$ 是 Lipschitz 连续的 (定理 6.1-1). 所以非线性复合映射 $f \in L^2(\Omega) \to u \in U$ 也如此.

(ii) 下面我们证明, 如果 $u \in H^2(\Omega)$, 则 $-\Delta u = f$ 在 $L^2(\Omega^+)$ 中成立, 其中 Ω^+ 由 $\Omega^+ := \{y \in \Omega; u(y) > \chi(y)\}$ 定义.

给定任意点 $x \in \Omega^+$, 令 $2\delta := u(x) - \chi(x) > 0$. 因 $x \in \Omega$ 且 Ω 是开的, 又因函数 $(u-\chi) : \Omega \to \mathbb{R}$ 是连续的, 故存在 $r > 0$ 使得

$$B(x;r) \subset \Omega \text{ 且 } u(y) - \chi(y) \geqslant \delta \quad \text{对所有 } y \in B(x;r),$$

这证明集合 $B(x,r) \subset \Omega^+$, 因此 Ω^+ 是开的.

给定任意非零函数 $\varphi \in \mathcal{D}(\Omega)$ 使得 $\operatorname{supp} \varphi \subset B(x;r)$, 令

$$\alpha_0 = \alpha_0(\varphi) := \frac{\delta}{\sup\limits_{y \in B(x;r)} |\varphi(y)|} > 0.$$

因此函数 $v_\alpha := u + \alpha\varphi$ 属于集合 U 对所有 $|\alpha| \leqslant \alpha_0$, 这是因为

$$\begin{aligned} v_\alpha - \chi(y) &= u(y) - \chi(y) + \alpha\varphi(y) \\ &\geqslant \delta - |\alpha\varphi(y)| \geqslant 0 \quad \text{对所有 } y \in B(x,r) \text{ 及所有 } |\alpha| \leqslant \alpha_0, \\ v_\alpha(y) - \chi(y) &= u(y) - \chi(y) \geqslant 0 \quad \text{对所有 } y \in (\Omega - B(x;r)). \end{aligned}$$

借助于定理 6.7-1(a) 中的 Green 公式以及关系式 $v - u = 0$ 在 Γ 上, 变分不等式 $a(u, v-u) \geqslant \ell(v-u)$ 对所有 $v \in U$ 就化为

$$\begin{aligned} \int_\Omega \nabla u \cdot \nabla(v-u) dx &= -\int_\Omega \Delta u (v-u) dx \\ &\geqslant \int_\Omega f(v-u) dx \quad \text{对所有 } v \in U. \end{aligned}$$

在上述不等式中令 $v = v_\alpha, |\alpha| \leqslant \alpha_0$, 就得到

$$\alpha \int_{B(x;r)} (-\Delta u - f)\varphi dx \geqslant 0 \quad \text{对所有 } \varphi \in \mathcal{D}(B(x;r)) \text{ 及所有 } |\alpha| \leqslant \alpha_0.$$

这又意味着 $\int_{B(x:r)} (-\Delta u - f)\varphi dx = 0$ 对所有 $\varphi \in \mathcal{D}(B(x;r))$ 成立. 于是 $-\Delta u = f$ 在 $L^2(B(x;r))$ 中, 所以 $-\Delta u = f$ 在 $L^2(\Omega^+)$ 中成立.

(iii) 剩下来的就是要证明, 如果还有 $u \in H^2(\Omega)$, 则 $-\Delta u - f \geqslant 0$ 几乎处处在 $\Omega^0 := \{y \in \Omega; u(y) = \chi(y)\}$ 中成立.

给定任意函数 $\varphi \in \mathcal{D}(\Omega)$, 其满足 $\varphi \geqslant 0$ 在 Ω 内, 则函数 $v := u + \varphi$ 属于 U. 因此, 对这样的函数 v, 像上面那样, 变分不等式结合与上面同样的 Green 公式就意味着

$$\int_\Omega (-\Delta u - f)(v-u) dx = \int_\Omega (-\Delta u - f)\varphi dx \geqslant 0$$
$$\text{对所有 } \varphi \in \mathcal{D}(\Omega) \text{ 且 } \varphi \geqslant 0 \text{ 在 } \Omega \text{ 内}.$$

但是, 如果函数 $w \in L^1(\Omega)$ 满足 $\int_\Omega w\varphi dx \geqslant 0$ 对所有 $\varphi \in \mathcal{D}(\Omega)$ 且 $\varphi \geqslant 0$ 在 Ω 内, 则 $w \geqslant 0$ 在 Ω 内几乎处处成立 (习题 2.6-5). 所以函数 $(-\Delta u - f) \in L^2(\Omega)$ 满足 $-\Delta u - f \geqslant 0$ 在 Ω 内几乎处处成立, 当然在 Ω^0 中也如此 (实际上, 由 (ii) 我们得 $-\Delta u - f = 0$ 在 Ω^+ 中几乎处处成立). □

下面应当给出几点说明. 第一, 定理 6.9-1 中所考虑的问题提供一个在映射 $f \in L^2(\Omega) \to u \in U$ 是非线性的意义下非线性问题的实例. 如同至今所讨论过的线性问题, 它也是适定的, 因为映射 $f \in L^2(\Omega) \to u \in U$ 是连续的.

第二, 对线性薄膜问题的解 (6.7 节) 可要求它具有我们所满意的光滑性, 但与此不同的是, 障碍问题的解一般是不光滑的, 即使资料是很光滑的. 为说明情况确实如此, 我们考察出现在定理 6.9-1 中的边值问题的一维类似模型, 其中 $f = 0$. 如图 6.9-2 所示, 在未触碰到障碍的区域解 u 是仿射的, 因此不管函数 χ 多么光滑, u 的二阶导数在如 ξ 和 η 这种点上是间断的. 所以解 u "只是" 在空间 $H^2(I)$ 中, 即使在这种简单的情况.

把这种考察进行到二维的情况, 如我们所预期的, 要验证这一点就没那么容易了. 例如, 现在已经知道, 如果 $f = 0$, $\chi \in H^2(\Omega)$ 且 $\overline{\Omega}$ 为一凸多边形, 则解 u 属于空间 $H_0^1(\Omega) \cap H^2(\Omega)$; 或如果 $\overline{\Omega}$ 是凸的, 其边界是 $\mathcal{C}^2$ 类, 则也有 $u \in H_0^1(\Omega) \cap H^2(\Omega)$. 此外, 在上述两种情况中, 范数 $\|u\|_{2,\Omega}$ 可用数据的范数 $\|\chi\|_{2,\Omega}$ 及 $\|f\|_{0,\Omega}$ 估计[37].

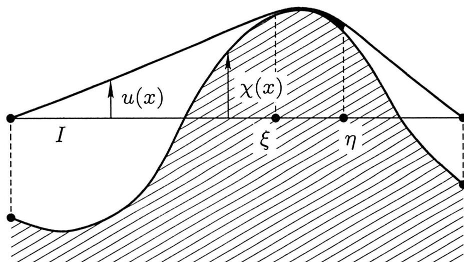


图 6.9-2 在 $\mathbb{R}$ 的有界开区间 I 上障碍问题的一维类似模型, 其中 $f = 0$. 该图最早出现在下文中: P. G. Ciarlet [1978]: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam

第三, 薄膜触及障碍的区域, 即集合 Ω^0 是事先不知道的.

第四, 上面的边值问题也可视为自由边界问题的一个例子. 在这种意义下, "自由边界" $\Gamma^* := \partial\Omega^+ \cap \partial\Omega^0$ 是问题的未知量之一. 以这种看法处理时, 通常沿边值问题未知的自由边界需要两个传递条件相连接, 即

$$\operatorname{tr}(u|_{\Omega^+}) = \operatorname{tr}(u|_{\Omega^0}) \quad \text{及} \quad \operatorname{tr} \partial_\nu(u|_{\Omega^+}) = -\operatorname{tr} \partial_\nu(u|_{\Omega^0}) \text{ 在 } \Gamma^* \text{ 上}.$$

[37] H. Brezis; G. Stampacchia [1968]: Sur la régularité de la solution d'inéquations elliptiques. Bulletin de la Société Mathématique de France **96**, 153–180.

H. Lewy; G. Stampacchia [1969]: On the regularity of the solution of a variational inequality. Communications on Pure and Applied Mathematics **22**, 153–188.

这些及其他类似结果的证明, 见 Kinderlehrer & Stampacchia [1980].

它们只有当 Γ^* 足够光滑 (例如, Γ^* 是一区域的边界) 且 u 也足够光滑 (例如, $u \in H^2(\Omega)$) 时才有意义.

另一些与变分不等式相应的边值问题实例, 包括板的障碍问题, 在习题 6.9-1 到 6.9-3 中给出.

习题

6.9-1 设 Ω 为 $\mathbb{R}^N$ 中的区域, 函数

$$c \in L^\infty(\Omega) \text{ 且 } c \geqslant c_0 > 0 \text{ 几乎处处在 } \Omega \text{ 内}, \quad f \in L^2(\Omega), g \in L^2(\Gamma)$$

是给定的, 又设

$$\begin{aligned}
&V := H^1(\Omega) \quad \text{及} \quad U := \{v \in H^1(\Omega); v \geqslant 0 \ d\Gamma \text{ 几乎处处在 } \Gamma \text{ 上}\},\\
&a(u,v) := \int_\Omega (\nabla u \cdot \nabla v + cuv)dx\\
&\text{及} \quad \ell(v) = \int_\Omega fvdx + \int_\Gamma gvdx \quad \text{对所有 } u, v \in H^1(\Omega).
\end{aligned}$$

(1) 证明相应的变分不等式有唯一解 $u \in U$.

(2) 证明, 如果 $u \in H^2(\Omega)$[38], 则 u 满足非线性边值问题:

$$\begin{aligned}
&-\Delta u + cu = f \quad \text{在 } \Omega \text{ 内},\\
&u \geqslant 0 \ d\Gamma \text{ 几乎处处在 } \Gamma \text{ 上}, \quad \partial_\nu u \geqslant g \ d\Gamma \text{ 几乎处处在 } \Gamma \text{ 上},\\
&\text{及 } u(\partial_\nu u - g) = 0 \ d\Gamma \text{ 几乎处处在 } \Gamma \text{ 上}.
\end{aligned}$$

注 这种所有边界条件或部分边界条件取不等式形式的边值问题称为 Signorini 问题[39].

6.9-2 下面的变分问题还可作为细圆柱形线性弹性杆的弹 – 塑性扭转的模型. 设 Ω 是 $\mathbb{R}^2$ 中的区域, 令

$$\begin{aligned}
&V := H_0^1(\Omega), \quad U := \{v \in H_0^1(\Omega); |\nabla v| \leqslant 1 \text{ 几乎处处在 } \Omega \text{ 内}\},\\
&a(u,v) := \int_\Omega \nabla u \cdot \nabla v dx \text{ 及 } \ell(v) := \tau \int_\Omega v dx,
\end{aligned}$$

---

[38] 如果 $g = 0$ 及 $N = 2$, 而又如果 Γ 是足够光滑的或者 Ω 是凸的而 Γ 是多边形, 则这个正则性假设满足. 见:

H. BREZIS [1971]: Problèmes unilatéraux. Journal de Mathématiques Pures et Appliquées **9**, 1–168.

[39] 冠名源自:

A. SIGNORINI: Sopra alcune questioni di elastostatica. Atti della Società Italiana per il Progresso della Scienza (1933).

Signorini 问题的第一个数学分析属于:

G. FICHERA [1964]: Problemi elastostatici con vincoli unilaterali: il problema de Signorini con ambigue condizioni al contorno. Memorie dell'Accademia Nazionale dei Lincei **8**, 91–140.

自此以后, 出现大量关于 Signorini 问题的研究工作, 重要的可见 FICHERA [1972b], DUVAUT & LIONS [1972], 以及 NEČAS & HLAVÁČEK [1981].

其中常数 $\tau \in \mathbb{R}$ 是测度杆扭转的量[40].

(1) 证明相应的变分不等式有唯一解 $u_\tau \in U$.

(2) 证明, 如果 $u_\tau \in H^2(\Omega)$, 则 u_τ 满足

$$-\Delta u_\tau = \tau \quad \text{几乎处处在集合 } \{x \in \Omega; |\nabla v(x)| < 1\} \text{ 内}.$$

(3) 证明集合 U 是 $\mathcal{C}(\overline{\Omega})$ 的紧子集而且任一函数 $v \in U$ 都满足 $|v(x)| \leqslant \operatorname{dist}(x, \Gamma)$ 对所有 $x \in \overline{\Omega}$.

(4) 证明, 当 $\tau \to \infty$ 时, $\|u_\tau - u_\infty\|_{1,\Omega} \to 0$ 及 $\sup_{x\in\overline{\Omega}} |u_\tau(x) - u_\infty(x)| \to 0$, 其中函数 $u_\infty : \overline{\Omega} \to \mathbb{R}$ 由 $u_\infty(x) := \operatorname{dist}(x, \partial\Omega)$ 对所有 $x \in \overline{\Omega}$ 定义[41].

6.9-3 设 Ω 为 $\mathbb{R}^N$ 中的区域, 其中 $N = 2$ 或 3, 又设 $x_i, 1 \leqslant i \leqslant m$, 为 Ω 中不同的点, $f \in L^2(\Omega)$, 令

$$V := H_0^2(\Omega), \quad U := \{v \in H_0^2(\Omega); v(x_i) \geqslant 0, 1 \leqslant i \leqslant m\},$$
$$a(u, v) := \int_\Omega \Delta u \Delta v dx \text{ 及 } \ell(v) := \int_\Omega f v dx \quad \text{对所有 } u, v \in H_0^2(\Omega).$$

(1) 证明相应的变分不等式有唯一解 $u \in U$.

(2) 证明, 如果 $u \in H^4(\Omega)$, 则 u 满足 $\Delta^2 u = f$ 在集合 $\Omega - \bigcup_{i=1}^m \{x_i\}$ 内.

注 (1) 如果 $N = 2$, 由 $J(v) = \dfrac{1}{2}\displaystyle\int_\Omega |\Delta v|^2 dx - \int_\Omega f v dx$ 对所有 $v \in H_0^2(\Omega)$, 定义的泛函 $J : H_0^2(\Omega) \to \mathbb{R}$ 表示沿其整个边界固定的线性弹性板的能量 (6.8 节). 因此上述变分问题就是一个*边界固定板的障碍问题的模型*, 其中未知的垂直位移 $u : \overline{\Omega} \to \mathbb{R}$ 受制于不等式 $u(x_i) \geqslant 0, 1 \leqslant i \leqslant m$.

(2) 问题 (2) 的有趣的补充将在习题 7.15-4 中给出, 在那里将证明, 存在 "Kuhn-Tucker 乘子" $\lambda_i \geqslant 0, 1 \leqslant i \leqslant m$, 满足 $\Delta^2 u = f + \sum_{i=1}^m \lambda_i \delta_{x_i}$ 在 $\mathcal{D}'(\Omega)$ 中 (即在分布的意义下, 见 6.3 节), 并且有值得关注的力学诠释.

6.10 二阶椭圆算子的特征值问题

设 Ω 是 $\mathbb{R}^N$ 中的区域, $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ 是形如定理 6.7-6 中考虑的连续且 $H_0^1(\Omega)$ 强制对称双线性形式, 而 $f \in L^2(\Omega)$. 因此存在唯一的函数 $u \in H_0^1(\Omega)$ 满足

$$a(u, v) = \langle f, v \rangle \quad \text{对所有 } v \in H_0^1(\Omega),$$

---

[40] 首先讨论这种变分问题的是:

H. Brezis; M. Sibony [1971]: Equivalence de deux inéquations variationnelles. Archive for Rational Mechanics and Analysis **41**, 254–265.

R. Glowinski; H. Lanchon [1973]: Torsion élasto-plastique d'une barre cylindrique de section multiconnexed. Journal de Mécanique **12**, 151–171.

更一般的弹 - 塑性问题在 Duvaut & Lions [1976] 及 Nečas & Hlaváček [1981] 中有详尽的讨论.

[41] 这一结果的证明在 Glowinski [1984, 第 2 章, 第 3 节].

其中 (整个这一节, 为符号简洁起见) 我们记

$$\langle f,g\rangle:=\int_\Omega fgdx \quad 对所有\ f,g\in L^2(\Omega).$$

此外, 这些方程的足够光滑的解满足形如

$$\mathcal{L}u=f\ 在\ \Omega\ 内,\quad u=0\ 在\ \Gamma:=\partial\Omega\ 上$$

的二阶椭圆边值问题, 其中 $\mathcal{L}$ 是一致椭圆型二阶线性偏微分算子 (6.7 节).

算子 $\mathcal{L}$ 的特征值问题是指, 寻求是否存在实数 μ 及非零函数 w 满足边值问题

$$\mathcal{L}w=\mu w\ 在\ \Omega\ 中,\quad w=0\ 在\ \Gamma\ 上.$$

如果偶对 (μ,w) 存在, μ 称为 $\mathcal{L}$ 的特征值而 w 称为 $\mathcal{L}$ 的相应特征值 μ 的特征函数 (当然, 每个这种特征函数都应该足够光滑使得上述边值问题有意义). 如果 $w\in H_0^1(\Omega)$, 偶对 $(\mu,w)\in\mathbb{R}\times H_0^1(\Omega)$ 就满足变分方程

$$a(w,v)=\mu\langle w,v\rangle \quad 对所有\ v\in H_0^1(\Omega),$$

上式构成算子 $\mathcal{L}$ 的特征值问题的变分形式.

只看上式本身, 即不考虑其与椭圆算子 $\mathcal{L}$ 的特征值问题的关系, 这个变分方程给出了抽象变分问题的另一个例子.

我们现在证明求解这些变分方程等价于求作用于 Hilbert 空间 $H_0^1(\Omega)$ 的紧对称正定算子特征值的倒数及相应的特征向量, 而空间 $H_0^1(\Omega)$ 中装备内积 $a(\cdot,\cdot)$.

要注意, 下面证明中虽然也用同一个符号 A, 但它与在 Lax-Milgram 引理 (定理 6.2-1) 证明中引入的算子不是同一个算子.

定理 6.10-1　设 Ω 是 $\mathbb{R}^N$ 中的区域, $a:H_0^1(\Omega)\times H_0^1(\Omega)\to\mathbb{R}$ 是连续且 $H_0^1(\Omega)$ 强制的对称双线性形式. 给定任意函数 $u\in H_0^1(\Omega)$, 于是就存在唯一的函数 $Au\in H_0^1(\Omega)$ 满足

$$a(Au,v)=\langle u,v\rangle \quad 对所有\ v\in H_0^1(\Omega).$$

(a) 以这种方式定义的线性算子 $A:H_0^1(\Omega)\to H_0^1(\Omega)$ 在 Hilbert 空间 $(H_0^1(\Omega),a(\cdot,\cdot))$ 中是紧的、对称且正定的, 因此是单射; 换言之

$$a(Au,v)=a(u,Av) \quad 对所有\ u,v\in H_0^1(\Omega),$$
$$a(Av,v)>0 \quad 对所有\ v\in H_0^1(\Omega),v\neq 0.$$

最后, A 有无限维的值域.

(b) 偶对 $(\mu,w)\in\mathbb{R}\times H_0^1(\Omega)$, 其中 $w\neq 0$, 满足

$$a(w,v)=\mu\langle w,v\rangle \quad 对所有\ v\in H_0^1(\Omega)$$

当且仅当

$$\mu > 0 \text{ 且 } Aw = \lambda w, \quad \text{其中 } \lambda := \frac{1}{\mu}.$$

证明　已经指出过 (见定理 6.1-1 的证明), 双线性形式 a 是空间 $H_0^1(\Omega)$ 上的一个内积, 其相应的范数等价于 $\|\cdot\|_{1,\Omega}$.

映射 $A: H_0^1(\Omega) \to H_0^1(\Omega)$ 显然是线性的. 此外, 映射

$$A: (H_0^1(\Omega), \|\cdot\|_{0,\Omega}) \to (H_0^1(\Omega), \|\cdot\|_{1,\Omega})$$

是连续的, 这是因为存在 $\alpha > 0$ 使得

$$\begin{aligned}\alpha\|Au\|_{1,\Omega}^2 &\leqslant a(Au, Au) = \langle u, Au\rangle \\ &\leqslant \|u\|_{0,\Omega}\|Au\|_{0,\Omega} \leqslant \|u\|_{0,\Omega}\|Au\|_{1,\Omega}\end{aligned}$$

对所有 $u \in H_0^1(\Omega)$.

设 $(u_n)_{n=1}^{\infty}$ 是 $H_0^1(\Omega)$ 中的有界序列; 由 Rellich-Kondrachov 定理 (定理 6.6-3), 存在一个子列 $(u_{\sigma(n)})_{n=1}^{\infty}$ 在 $L^2(\Omega)$ 中收敛. 映射 $A: (H_0^1(\Omega), \|\cdot\|_{0,\Omega}) \to (H_0^1(\Omega), \|\cdot\|_{1,\Omega})$ 的连续性意味着子列 $(Au_{\sigma(n)})_{n=1}^{\infty}$ 在 $H_0^1(\Omega)$ 中收敛. 所以 A 是紧的 (定理 2.10-1).

算子 A 在 Hilbert 空间 $(H_0^1(\Omega), a(\cdot,\cdot))$ 中的对称性及正定性可由下述关系式得到

$$\begin{aligned}&a(Au, v) = \langle u, v\rangle = \langle v, u\rangle = a(Av, u) = a(u, Av) \quad \text{对所有 } u, v \in H_0^1(\Omega), \\ &a(Av, v) = \langle v, v\rangle = \|v\|_{0,\Omega}^2 > 0 \quad \text{对所有 } v \in H_0^1(\Omega), v \neq 0.\end{aligned}$$

最后一个关系式也证明了, $Av = 0$ 意味着 $v = 0$.

于是 $A: H_0^1(\Omega) \to H_0^1(\Omega)$ 是单射, 又因为空间 $H_0^1(\Omega)$ 是无穷维的, 所以其值域也是无穷维的. (a) 中的断言均得证.

如果 $(\mu, w) \in \mathbb{R} \times H_0^1(\Omega)$, 其中 $w \neq 0$, 满足 $a(w, v) = \mu\langle w, v\rangle$ 对所有 $v \in H_0^1(\Omega)$, 那么特别地有 $\mu\langle w, w\rangle = a(w, w) > 0$, 这意味着 $\mu > 0$.

此外, 由 A 的定义

$$a(w, v) = \mu\langle w, v\rangle = \mu a(Aw, v) \quad \text{对所有 } v \in H_0^1(\Omega),$$

故 $Aw = \lambda w$, 其中 $\lambda := \frac{1}{\mu}$. 反之, 若 $(\mu, w) \in \mathbb{R} \times H_0^1(\Omega)$, 其中 $\mu \neq 0$ 及 $w \neq 0$, 满足 $Aw = \frac{1}{\mu}w$, 那么仍由 A 的定义

$$\mu\langle w, v\rangle = \mu a(Aw, v) = a(w, v) \quad \text{对所有 } v \in H_0^1(\Omega).$$

所以 (b) 得证. □

注　利用同一个关系式 $a(\widetilde{A}f, v) = \langle f, v\rangle$ 对所有 $v \in H_0^1(\Omega)$, 也可以定义另一个紧的、对称且正定的算子 $\tilde{A}$, 但此时它是映 Hilbert 空间 $L^2(\Omega)$ 至其本身; 见习题 6.10-2.

将定理 6.10-1 与紧对称算子的谱定理 (4.11 节) 结合起来, 立即可以看出, 该定理对于很广的一类一致椭圆型算子 $\mathcal{L}$ 给出了特征值问题: $\mathcal{L}w = \mu w$ 在 Ω 内, $w = 0$ 在 Γ 上 (本节开始所考虑的), 所有的解 (μ, w). 此外, 对于 $\mathcal{L}$ 的特征值及特征函数, 可以利用一个特定的泛函, 即下述定理中引入的 Rayleigh 商来刻画.

定理 6.10-2 设 Ω 是 $\mathbb{R}^N$ 中的区域, 函数 $a_{ij} = a_{ji} \in \mathcal{C}^1(\overline{\Omega}), 1 \leqslant i, j \leqslant N$, 是给定的使得对某常数 $\mu > 0$,

$$\sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geqslant \mu \sum_{i=1}^{N} |\xi_i|^2 \quad \text{对所有 } x \in \overline{\Omega} \text{ 及所有 } (\xi_i)_{i=1}^N \in \mathbb{R}^N,$$

又设函数 $c \in L^\infty(\Omega)$ 是给定的使得 $c \geqslant 0$ 在 Ω 内几乎处处成立. 由下式定义双线性形式 $a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$:

$$a(u, v) := \int_\Omega \left(\sum_{i,j=1}^{N} a_{ij}\partial_i u \partial_j v + cuv \right) dx \quad \text{对所有 } u, v \in H_0^1(\Omega).$$

(a) 存在一个实数无穷序列 $(\mu_k)_{k=1}^\infty$ 以及一个非零函数 $w_k \in H_0^1(\Omega)$ 的无穷序列 $(w_k)_{k=1}^\infty$ 满足

$$\begin{aligned}
&0 < \mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_k \leqslant \cdots, \quad \lim_{k\to\infty} \mu_k = \infty, \\
&a(w_k, v) = \mu_k \langle w_k, v\rangle \quad \text{对所有 } v \in H_0^1(\Omega) \text{ 及所有 } k \geqslant 1, \\
&a(w_k, w_\ell) = \delta_{k\ell} \quad \text{及} \quad \langle w_k, w_\ell\rangle = \frac{\delta_{k\ell}}{\mu_k} \quad \text{对所有 } k, \ell \geqslant 1.
\end{aligned}$$

设 $\mu \in \mathbb{R}$, 而非零函数 $w \in H_0^1(\Omega)$ 是以下变分方程的解:

$$a(w, v) = \mu\langle w, v\rangle \quad \text{对所有 } v \in H_0^1(\Omega).$$

则存在 $k \geqslant 1$ 使得 $\mu_k = \mu$. 此外, 集合 $J(\mu) := \{k \geqslant 1; \mu_k = \mu\}$ 是有限集, 且

$$\{w \in H_0^1(\Omega); a(w, v) = \mu\langle w, v\rangle \text{ 对所有 } v \in H_0^1(\Omega)\} = \operatorname{Span}(w_k)_{k\in J(\mu)}.$$

而且, 函数族 $(w_k)_{k=1}^\infty$ 是 Hilbert 空间 $(H_0^1(\Omega), a(\cdot,\cdot))$ 的 Hilbert 基 (4.9 节), 函数族 $(\sqrt{\mu_k}w_k)_{k=1}^\infty$ 是 Hilbert 空间 $(L^2(\Omega), \langle\cdot,\cdot\rangle)$ 的 Hilbert 基.

(b) 定义 Rayleigh 商[42)]

$$R(w) := \frac{a(w, w)}{\langle w, w\rangle} = \frac{\displaystyle\int_\Omega \left(\sum_{i,j=1}^{N} a_{ij}\partial_i w \partial_j w + c|w|^2 \right) dx}{\displaystyle\int_\Omega |w|^2 dx} \quad \text{对所有 } w \in H_0^1(\Omega), w \neq 0.$$

---

42) 冠名源自 John William Strut, Rayleigh 男爵三世 (1842—1919). Rayleigh 勋爵于 1904 年荣获 Nobel 物理学奖.

则

$$\mu_1 = R(w_1) = \inf_{w\in H_0^1(\Omega), w\neq 0} R(w),$$

$$\mu_k = R(w_k) = \inf_{\left\{\substack{w \in H_0^1(\Omega), w\neq 0 \\ \langle w, w_\ell\rangle = 0, 1\leqslant \ell \leqslant k-1}\right.} R(w) \quad \text{对所有 } k\geqslant 2.$$

(c) 如果 $w_k \in H^2(\Omega)$ 对某个 $k \geqslant 1$[43), 则

$$\mathcal{L}w_k = \mu_k w_k \text{ 在 } \Omega \text{ 内}, \quad w_k = 0 \text{ 在 } \Gamma \text{ 上},$$

其中一致椭圆算子 $\mathcal{L}$ 对所有足够光滑的函数 v 由下式定义

$$\mathcal{L}v := -\sum_{i,j=1}^{N} \partial_i(a_{ij}\partial_j v) + cv.$$

证明 前面考虑的将 Hilbert 空间 $(H_0^1(\Omega), a(\cdot,\cdot))$ 映射到其自身的线性算子 A 是紧的, 对称且正定的 (定理 6.10-1(a)). 所以由定理 4.11-1 及定理 4.11-3, 存在 A 的特征值无穷序列 $(\lambda_k)_{k=1}^\infty$ 及相应的特征向量无穷序列 $(w_k)_{k=1}^\infty$ 满足

$$\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_k \geqslant \cdots, \lambda_k > 0 \text{ 对所有 } k \geqslant 1, \lim_{k\to\infty} \lambda_k = 0,$$

$$Aw_k = \lambda_k w_k \quad \text{对所有 } k \geqslant 1 \quad \text{且} \quad a(w_k, w_\ell) = \delta_{k\ell} \quad \text{对所有 } k,\ell \geqslant 1,$$

$$\lambda_1 = \frac{a(Aw_1, w_1)}{a(w_1, w_1)} = \sup_{w\in H_0^1(\Omega), w\neq 0} \frac{a(Aw, w)}{a(w, w)},$$

$$\lambda_k = \frac{a(Aw_k, w_k)}{a(w_k, w_k)} = \sup_{\left\{\substack{w \in H_0^1(\Omega), \quad w\neq 0 \\ a(w, w_\ell) = 0, \quad 1\leqslant \ell \leqslant k-1}\right.} \frac{a(Aw, w)}{a(w, w)} \quad \text{对所有 } k \geqslant 2.$$

此外, 函数族 $(w_k)_{k=1}^\infty$ 是 Hilbert 空间 $(H_0^1(\Omega), a(\cdot,\cdot))$ 的 Hilbert 基.

关系式 $Aw_k = \lambda_k w_k$ 及 $a(w_k, w_\ell) = \delta_{k\ell}$ 意味着由定理 6.10-1(b) 有

$$\langle w_k, w_\ell\rangle = \lambda_k a(w_k, w_\ell) = \lambda_k \delta_{k\ell} \quad \text{对所有 } k,\ell \geqslant 1.$$

因此函数族 $(\lambda_k^{-\frac{1}{2}} w_k)_{k=1}^\infty$ 关于 $\langle\cdot,\cdot\rangle$ 是正交的, 为了证明它是空间 $(L^2(\Omega), \langle\cdot,\cdot\rangle)$ 的 Hilbert 基, 只需证明 (定理 4.8-2)

$$\overline{\mathrm{Span}(\lambda_k^{-\frac{1}{2}} w_k)_{k=1}^\infty} = \overline{\mathrm{Span}(w_k)_{k=1}^\infty} = L^2(\Omega),$$

其中两个闭包都是关于范数 $\|\cdot\|_{0,\Omega}$ 的.

为此, 设函数 $u \in L^2(\Omega)$ 及 $\varepsilon > 0$ 是给定的. 因为空间 $\mathcal{D}(\Omega)$ 在 $L^2(\Omega)$ 中稠密, 存在 $\varphi = \varphi(u,\varepsilon) \in \mathcal{D}(\Omega) \subset H_0^1(\Omega)$ 使得 $\|\varphi - u\|_{0,\Omega} \leqslant \dfrac{\varepsilon}{2}$. 又因为 $(w_k)_{k=1}^\infty$

[43) 如果 Γ 是 $\mathcal{C}^2$ 类的, 对所有 $k \geqslant 1$ 均是这种情况; 参见如 Evans [2010, 6.3 节].

是 $(H_0^1(\Omega); a(\cdot,\cdot))$ 的 Hilbert 基, 存在 $v = v(\varphi) = v(u,\varepsilon) \in \mathrm{Span}(w_k)_{k=1}^{\infty}$ 使得 $\|v - \varphi\|_{0,\Omega} \leqslant \|v - \varphi\|_{1,\Omega} \leqslant \dfrac{\varepsilon}{2}$. 所以 $\|v - u\|_{0,\Omega} \leqslant \varepsilon$, 这就明了 $\overline{\mathrm{Span}(w_k)_{k=1}^{\infty}} = L^2(\Omega)$.

取 $\mu_k := \dfrac{1}{\lambda_k}, k \geqslant 1$, (a) 的所有断言均已得证.

为证明 (b) 的断言, 我们首先注意, 算子 A 的定义说明, Rayleigh 商也可由下式给出

$$R(w) := \frac{a(w,w)}{\langle w,w\rangle} = \frac{a(w,w)}{a(Aw,w)}, \quad \text{对所有非零 } w \in H_0^1(\Omega).$$

其次, 我们要注意, 一个函数 $w \in H_0^1(\Omega)$ 满足 $a(w,w_\ell) = 0$ 对所有 $1 \leqslant \ell \leqslant k-1$ 的充分必要条件是 $\langle w, w_\ell\rangle = 0$ 对所有 $1 \leqslant \ell \leqslant k-1$, 这是由于

$$\langle w, w_\ell\rangle = a(Aw_\ell, w) = \lambda_\ell a(w, w_\ell) \text{ 且 } \lambda_\ell > 0, \quad \text{对所有 } \ell \geqslant 1.$$

所有由数 $\lambda_k = \dfrac{1}{\mu_k}$ 作为上确界的表示立即就得到 $\mu_k, k \geqslant 1$, 作为下确界的表示.

断言 (c) 可以利用 Green 公式以通常的方式 (见定理 6.7-6 的证明) 予以证明. □

定理 6.10-2 的作用从下面最简单的特征值问题例子已可见一斑:

$$-u''(x) = \mu u(x), \quad 0 < x < 1 \quad \text{且} \quad u(0) = u(1) = 0.$$

在这种情况下, 特征值 μ_k 相应的特征向量 $w_k, k \geqslant 1$, 关于由 $a(u,v) := \int_0^1 u'v'dx$ 定义的内积 $a(\cdot,\cdot)$ 正交. 它们由下式给出

$$\mu_k = k^2\pi^2, \quad w_k(x) = \frac{\sqrt{2}}{k\pi}\sin k\pi x, \quad 0 \leqslant x \leqslant 1.$$

这样, 由定理 6.10-2(a) 立即可得, 函数族 $(w_k)_{k=1}^{\infty}$ 构成 Hilbert 空间 $(H_0^1(0,1); a(\cdot,\cdot))$ 的 Hilbert 基, 而函数族 $(k\pi w_k)_{k=1}^{\infty}$ 构成空间 $L^2(0,1)$ 的 Hilbert 基. 由定理 6.10-2(b) 又立即可得有趣的公式

$$\pi^2 = \inf_{\left\{\substack{w \in H_0^1(0,1) \\ w \neq 0}\right.} \frac{\int_0^1 |w'|^2 dx}{\int_0^1 |w|^2 dx}.$$

注 容易验证, 在取下确界的式子中, 空间 $H_0^1(0,1)$ 可换为满足 $\mathcal{D}(0,1) \subset V \subset H_0^1(0,1)$ 的任一函数空间 V.

给定空间 $L^2(\Omega)$ 的一个子空间 W, 以 $W^\perp$ 表示其关于 $L^2(\Omega)$ 的内积 $\langle\cdot,\cdot\rangle$ 的正交补 (4.5 节). 这样, 在定理 6.10-2(b) 中给出的用 Rayleigh 商表示第 k 个特征值 μ_k 的式子可重写为

$$\mu_k = R(w_k) = \inf\{R(w); w \in W_{k-1}^\perp, w \neq 0\},$$

其中

$$W_0 := \{0\}, \quad W_{k-1} := \mathrm{Span}(w_\ell)_{\ell=1}^{k-1}, \quad \text{若 } k \geqslant 2.$$

值得关注的是, 特征值 $\mu_k, k \geqslant 1$, 还可以用 Rayleigh 商, 但是与特征函数无关的方式来刻画.

定理 6.10-3 (Courant-Fischer 定理[44]) 此定理的假设与定理 6.10-2 中的相同. 对每个整数 $\ell \geqslant 1$, 设 $\mathcal{V}_\ell$ 表示 $H_0^1(\Omega)$ 的所有 ℓ 维子空间的集合, 令 $\mathcal{V}_0 = \{0\}$. 则对每个整数 $k \geqslant 1$,

$$\mu_k = \sup_{V \in \mathcal{V}_{k-1}} (\inf\{R(w); w \in V^\perp, w \neq 0\}),$$
$$\mu_k = \inf_{V \in \mathcal{V}_k} (\sup\{R(w); w \in V, w \neq 0\}).$$

证明　为简洁起见, 在整个证明中将略去关系式 "$w \neq 0$".

假定 $k \geqslant 2$ (对于 $k = 1$, 第一个关系式显然成立, 由于如果 $V \in \mathcal{V}_0$, 即 $V = \{0\}$, 则 $V^\perp = H_0^1(\Omega)$). 因为 $W_{k-1} = \mathrm{Span}(w_\ell)_{\ell=1}^{k-1} \in \mathcal{V}_{k-1}$ (因为是正交的, 特征函数 $w_\ell, 1 \leqslant \ell \leqslant k-1$, 是线性无关的), 就有

$$\mu_k = \inf\{R(w); w \in W_{k-1}^\perp\} \leqslant \sup_{V \in \mathcal{V}_{k-1}} (\inf\{R(w); w \in V^\perp\}).$$

剩下来要证明, 给定任意子空间 $V \in \mathcal{V}_{k-1}$,

$$\inf\{R(w); w \in V^\perp\} \leqslant \mu_k.$$

为此目的, 我们注意, 存在函数 u 满足

$$u \in \mathrm{Span}(w_j)_{j=1}^k, \quad u \neq 0 \quad \text{且 } u \in V^\perp,$$

这是因为, 给定 V 中的基 $(v_i)_{i=1}^{k-1}$, 齐次线性方程组

$$\sum_{j=1}^k \alpha_j \langle w_j, v_j \rangle = 0, \quad 1 \leqslant i \leqslant k-1,$$

总有非零解. 给定这样一个非零解, 令 $u := \sum_{j=1}^k \alpha_j w_j$. 这样, 特征函数所满足的正交性关系给出

$$R(u) = R\left(\sum_{j=1}^k \alpha_j w_j\right) = \frac{\sum_{j=1}^k |\alpha_j|^2}{\sum_{j=1}^k \mu_j^{-1} |\alpha_j|^2} \leqslant \mu_k,$$

[44] 这个定理先是对于矩阵确立的, 后来才推广到我们这里所考虑的这类特征值问题, 见:

E. Fischer [1905]: Über quadratische Formen mit reellen Koeffzienten. Monatsheft für Mathematik und Physik **16**, 234–249.

R. Courant [1920]: Über die Eigenwerte bei den Differentialgleichungen der Mathematischen Physik. Mathematische Zeitschrift **7**, 1–57.

此因 $0 < \mu_1 \leqslant \cdots \leqslant \mu_k$. 所以 $\inf\{R(w); w \in V^\perp\} \leqslant \mu_k$, 故

$$\mu_k = \sup_{V \in \mathcal{V}_{k-1}} (\inf\{R(w); w \in V^\perp\}).$$

为证明另一个关系式, 我们首先注意

$$\sup_{w \in W_k = \mathrm{Span}(w_\ell)_{\ell=1}^k} R(w) = \sup_{(\alpha_j)_{j=1}^k \in \mathbb{R}^k - \{0\}} R\left(\sum_{j=1}^k \alpha_j w_j\right) = \mu_k = R(w_k) \quad \text{对所有 } k \geqslant 1,$$

这就有

$$\mu_k = \sup_{w \in W_k} R(w) \geqslant \inf_{V \in \mathcal{V}_k} (\sup\{R(w); w \in V\}),$$

此因 $W_k \in \mathcal{V}_k$. 剩下的就是要证明, 给定任意的子空间 $V \in \mathcal{V}_k$, 有

$$\mu_k \leqslant \sup\{R(w); w \in V\}.$$

为此, 我们注意, 一定存在函数 u 满足

$$u \in V, \quad u \neq 0, \quad u \in W_{k-1}^\perp.$$

这是因为, 给定 V 中关于 $\langle \cdot, \cdot \rangle$ 正交的基 $(v_j)_{j=1}^k$, 齐次线性方程组

$$\sum_{j=1}^k \beta_j \langle v_j, w_i \rangle = 0, \quad 1 \leqslant i \leqslant k-1,$$

总有满足 $\beta_k \neq 0$ 的解. 给定这样一个解, 令 $u := \sum_{j=1}^k \beta_j v_j$. 由于关系式 $\langle u, w_\ell \rangle = 0, 1 \leqslant \ell \leqslant k-1$, 意味着 $a(u, w_\ell) = 0, 1 \leqslant \ell \leqslant k-1$, 由定理 6.10-2(b) 得 $\mu_k = \inf\{R(w); w \in W_{k-1}^\perp\} \leqslant R(u)$. 所以

$$\mu_k = \inf_{V \in \mathcal{V}_k} (\sup\{R(w); w \in V\}). \qquad \square$$

注 (1) 在用有限元方法求解这里讨论的这类特征值问题时, Courant-Fischer 定理在其数值逼近的收敛分析中起着实质性的作用[45].

(2) Courant-Fischer 定理立即可被推广为一个对更一般的任意紧自伴算子 (4.10 节) 都成立的类似定理.

习题

6.10-1 直接计算下列特征值问题的所有特征值及相应的特征函数:

$$-u''(x) + u(x) = \mu u(x), \quad 0 < x < 1 \text{ 且 } u(0) = u(1) = 0,$$

$$-u''(x) + u(x) = \mu u(x), \quad 0 < x < 1 \text{ 且 } u'(0) = u'(1) = 0,$$

$$-u''(x) + u(x) = \mu u(x), \quad 0 < x < 1 \text{ 且 } u(0) = u(1) \text{ 及 } u'(0) = u'(1).$$

[45] I. Babuška; J.E. Osborn [1991]: Eigenvalue problems. Handbook of Numerical Analysis, Volume II (P.G. Ciarlet & J. L. Lions, editors), pp. 641–787, North-Holland, Amsterdam.

6.10-2 设 Ω 为 $\mathbb{R}^N$ 中的区域, 令 $a: H_0^1(\Omega)\times H_0^1(\Omega)\to\mathbb{R}$ 是连续且 $H_0^1(\Omega)$ 强制的对称双线性形式. 给定任意函数 $f\in L^2(\Omega)$, 存在唯一函数 $\widetilde{A}f\in H_0^1(\Omega)\subset L^2(\Omega)$ 满足

$$a(\widetilde{A}f,v)=\langle f,v\rangle \quad 对所有\ v\in H_0^1(\Omega).$$

证明, 以这种方式定义的线性算子 $\widetilde{A}: L^2(\Omega)\to L^2(\Omega)$ 在 Hilbert 空间 $L^2(\Omega)$ 中是紧的, 对称且正定的 (即 $\langle\widetilde{A}f,g\rangle=\langle f,\widetilde{A}g\rangle$ 对所有 $f,g\in L^2(\Omega)$ 及 $\langle\tilde{A}f,f\rangle>0$ 对所有 $f\in L^2(\Omega), f\neq 0$); 该算子还是单射, 且有无限维的值域.

6.10-3 设 Ω_1 与 Ω_2 为 $\mathbb{R}^N$ 中的两个区域, 且 $\Omega_1\subset\Omega_2$. 证明, 如在定理 6.10-2 中那样, 按递增顺序排列的算子 $-\Delta$ 相应的特征值 $\mu_k(\Omega_1)$ 及 $\mu_k(\Omega_2)$ 满足 $\mu_k(\Omega_2)\leqslant\mu_k(\Omega_1)$ 对所有 $k\geqslant 1$.

6.10-4 设 $V:=\left\{v\in H^1(-1,1); v(-1)=v(1)\ 且\ \int_{-1}^1 v(x)dx=0\right\}$. 证明[46]

$$\pi^2=\inf_{\left\{\substack{v\in V\\ v\neq 0}\right.}\frac{\int_{-1}^1|v'|^2dx}{\int_{-1}^1|v|^2dx}.$$

6.10-5 设 Ω 是 $\mathbb{R}^N$ 中的区域, $\mu_1>0$ 表示在 Ω 上的算子 $-\Delta$ 的最小特征值, 又设 $f:\mathbb{R}\to\mathbb{R}$ 是 Lipschitz 连续的函数, 其 Lipschitz 常数为 γ. 证明, 如果 $\gamma<\mu_1$, 则半线性边值问题

$$-\Delta u=f(u)\ 在\ \Omega\ 内, \quad u=0\ 在\ \Gamma\ 上,$$

在 $H_0^1(\Omega)$ 中有且只有一个解.

提示: 证明 u 是从 $L^2(\Omega)$ 到其自身, Lipschitz 常数为 $\gamma\mu_1^{-1}$ 的 Lipschitz 连续映射的不动点.

注 当 $N=1$ 而 $\Omega=]0,1[$ 时, 这个结果是定理 3.9-1 的一个改进 (Banach 不动点定理的另一个应用, 但是在不同的 Banach 空间里), 在那个定理中假设 $\gamma<8$, 这里 $\gamma<\pi^2$.

6.11 空间 $W^{-m,q}(\Omega)$ 与 $H^{-m}(\Omega)$; J. L. Lions 引理

空间 $W^{m,p}(\Omega)$ 与 $W_0^{m,p}(\Omega), 1\leqslant p<\infty$, 已在 6.5 节和 6.6 节中引入并予以讨论. 我们现在给出其对偶空间.

我们回忆一下, 对任意的 $1\leqslant p<\infty$, 其共轭指数 q 由下式定义: $q=\dfrac{p}{p-1}$ 如果 $1<p<\infty$, 而 $q:=\infty$ 如果 $p=1$. 给定一个整数 $m\geqslant 1$ 及对每一个重指标 $|\boldsymbol{\alpha}|\leqslant m$ 给定函数 $f_\alpha\in L^q(\Omega)$, 线性泛函 $v\in W^{m,p}(\Omega)\to\sum_{|\boldsymbol{\alpha}|\leqslant m}\int_\Omega f_{\boldsymbol{\alpha}}\partial^{\boldsymbol{\alpha}}vdx$ 显然在空间 $W^{m,p}$ 中连续. 我们现在证明, *反之*, $W^{m,p}(\Omega)$ 上的任一连续线性泛函一定具有这种形式. 然而要注意, 给定这样一个泛函, 出现在下述定理中的函数 $f_{\boldsymbol{\alpha}}\in L^q(\Omega)$, $|\boldsymbol{\alpha}|\leqslant m$, 并不一定是唯一的.

[46] 由此得到的不等式 $\int_{-1}^1|v'|^2dx\geqslant\pi^2\int_{-1}^1|v|^2dx$ 对所有 $v\in V$ 称为 Wirtinger 不等式, 冠名源自 Wilhelm Wirtinger (1865—1945).

定理 6.11-1 ($W^{m,p}(\Omega)$ 的对偶空间, $1 \leqslant p < \infty$) 设 Ω 是 $\mathbb{R}^N$ 的开子集 ,$m \geqslant 1$ 是整数, $1 \leqslant p < \infty$, 又令 q 表示 p 的共轭指数.

则 $\ell \in (W^{m,p}(\Omega))'$ 的充分必要条件是, 对每个重指标 $\boldsymbol{\alpha}, |\boldsymbol{\alpha}| \leqslant m$, 存在相应的函数 $f_{\boldsymbol{\alpha}} \in L^q(\Omega)$ 使得 (记 $\partial^0 v := v$)

$$\ell(v) = \sum_{|\boldsymbol{\alpha}| \leqslant m} \int_\Omega f_\alpha \partial^{\boldsymbol{\alpha}} v dx \quad \text{对所有 } v \in W^{m,p}(\Omega).$$

证明 容易看出, 充分性部分是显然的. 令 $M := \mathrm{Card}\{\boldsymbol{\alpha}; |\boldsymbol{\alpha}| \leqslant m\}$. 作为一个赋范向量空间, $W^{m,p}(\Omega)$ 可以等同于装备以范数 $(v^{\boldsymbol{\alpha}}) \to (\sum_{|\boldsymbol{\alpha}| \leqslant m} \|v^{\boldsymbol{\alpha}}\|_{0,p,\Omega}^p)^{1/p}$ 的积空间 $(L^p(\Omega))^M$ 的子空间.

$$Y(\Omega) := \left\{ (v^{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}| \leqslant m} \in (L^p(\Omega))^M; \int_\Omega v^{\boldsymbol{\alpha}} \varphi dx = (-1)^{|\boldsymbol{\alpha}|} \int_\Omega v^0 \partial^{\boldsymbol{\alpha}} \varphi dx \right.$$
$$\left. \text{对所有 } \varphi \in \mathcal{D}(\Omega), |\boldsymbol{\alpha}| \leqslant m \right\}$$

(显然子空间 $Y(\Omega)$ 在 $(L^p(\Omega))^M$ 中是闭的, 但这一性质在下面的证明中并不需要). 由赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1), 任何连续线性泛函 $\ell : W^{m,p}(\Omega) \to \mathbb{R}$ 可以延拓为连续线性泛函 $\widetilde{\ell} : (L^p(\Omega))^M \to \mathbb{R}$.

由 $L^p(\Omega), 1 \leqslant p < \infty$, 中的 F. Riesz 表示定理 (定理 3.5-3), 存在函数 $f_{\boldsymbol{\alpha}} \in L^q(\Omega)$ 使得

$$\widetilde{\ell}((v^{\boldsymbol{\alpha}})) = \sum_{|\boldsymbol{\alpha}| \leqslant m} \int_\Omega f_{\boldsymbol{\alpha}} v^{\boldsymbol{\alpha}} dx \quad \text{对所有 } (v^{\boldsymbol{\alpha}}) \in (L^p(\Omega))^M.$$

这样, 由 $\widetilde{\ell}$ 在 $(L^p(\Omega))^M$ 的子空间 $Y(\Omega)$ 上的限制, 即得必要性部分的证明. □

下面我们来确定 Sobolev 空间 $W_0^{m,p}(\Omega)$ 的对偶空间 $(W_0^{m,p}(\Omega))'$. 回忆一下, $\mathcal{D}(\Omega)$ 表示 Ω 上所有分布的空间 (6.3 节). 注意, 给定一个泛函 $\ell \in (W_0^{m,p}(\Omega))'$, 出现在下述定理中的函数 $f_{\boldsymbol{\alpha}} \in L^q(\Omega), |\boldsymbol{\alpha}| \leqslant m$, 仍不一定是唯一的.

定理 6.11-2 (对偶空间 $W_0^{m,p}(\Omega), 1 \leqslant p < \infty$) 设 Ω 是 $\mathbb{R}^N$ 的开子集, $m \geqslant 1$ 是整数, $1 \leqslant p < \infty$, 又令 q 表示 p 的共轭指数.

则对偶空间 $(W_0^{m,p}(\Omega))'$ 等同于所有形如

$$T = \sum_{|\boldsymbol{\alpha}| \leqslant m} (-1)^{|\boldsymbol{\alpha}|} \partial^\alpha f_{\boldsymbol{\alpha}}, \quad \text{对某些 } f_{\boldsymbol{\alpha}} \in L^q(\Omega), |\boldsymbol{\alpha}| \leqslant m,$$

的分布 $T \in \mathcal{D}'(\Omega)$ 组成的空间.

证明 由赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1), 任何连续线性泛函 $\ell : W_0^{m,p}(\Omega) \to \mathbb{R}$ 可以延拓为连续线性泛函 $\widehat{\ell} : W^{m,p}(\Omega) \to \mathbb{R}$. 由定理 6.11-1, 存在函数 $f_{\boldsymbol{\alpha}} \in L^q(\Omega)$ 使得

$$\widehat{\ell}(v) = \sum_{|\boldsymbol{\alpha}|\leqslant m} \int_\Omega f_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}} v dx, \quad \text{对所有 } v \in W^{m,p}(\Omega),$$

因此

$$\ell(v) = \sum_{|\alpha|\leqslant m} \int_\Omega f_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}} v dx, \quad \text{对所有 } v \in W_0^{m,p}(\Omega).$$

特别地, 由分布意义求导的定义有

$$\ell(\varphi) = \left(\sum_{|\boldsymbol{\alpha}|\leqslant m} (-1)^{|\boldsymbol{\alpha}|} \partial^{\boldsymbol{\alpha}} f_\alpha \right)(\varphi), \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega) \subset W_0^{m,p}(\Omega).$$

这说明, ℓ 在 $W_0^{m,p}(\Omega)$ 的子空间 $\mathcal{D}(\Omega)$ 上的限制是由下式

$$T := \sum_{|\boldsymbol{\alpha}|\leqslant m} (-1)^{|\boldsymbol{\alpha}|} \partial^{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}}$$

定义的分布 $T \in \mathcal{D}'(\Omega)$.

反之, 设给定这种形式的一个分布 $T \in \mathcal{D}'(\Omega)$, 其中 $f_{\boldsymbol{\alpha}} \in L^q(\Omega), |\boldsymbol{\alpha}| \leqslant m$. 则 T 是装备以范数 $\|\cdot\|_{m,p,\Omega}$ 的空间 $\mathcal{D}(\Omega)$ 上的连续线性泛函, 这是因为

$$\begin{aligned}|T(\varphi)| = \left| \left(\sum_{|\boldsymbol{\alpha}|\leqslant m} (-1)^{|\boldsymbol{\alpha}|} \partial^{\boldsymbol{\alpha}} f_\alpha \right)(\varphi) \right| &\leqslant \sum_{|\boldsymbol{\alpha}|\leqslant m} \left| \int_\Omega f_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}} \varphi dx \right| \\ &\leqslant \|(f_{\boldsymbol{\alpha}})_{|\boldsymbol{\alpha}|\leqslant m}\|_{(L^q(\Omega))^M} \|\varphi\|_{m,p,\Omega}, \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega),\end{aligned}$$

其中 $M := \mathrm{Card}\{\boldsymbol{\alpha}; |\boldsymbol{\alpha}| \leqslant m\}$. 因为根据定义, $\mathcal{D}(\Omega)$ 在空间 $W_0^{m,p}(\Omega)$ 中稠密, 故分布 T 在空间 $W_0^{m,p}(\Omega)$ 上有唯一的连续线性延拓. 这就证明了 $(W_0^{m,p}(\Omega))'$ 实际上可以等同于所有这种分布 T 组成的空间. □

设 Ω 是 $\mathbb{R}^N$ 的开子集. 对每一个整数 $m \geqslant 1$ 及每一个实数 $1 \leqslant p < \infty$, 在定理 6.11-2 中等同的对偶空间表示为

$$W^{-m,q}(\Omega) := (W_0^{m,p}(\Omega))',$$

其中 q 表示 p 的共轭指数, 或

$$H^{-m}(\Omega) := (H_0^m(\Omega))' \quad \text{如果 } p = 2.$$

例如, 给定任意整数 $m \geqslant 1$, 与一个函数相应的, 即由

$$T_v(\varphi) = \int_\Omega v\varphi dx, \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega)$$

定义的分布 T_v (6.3 节), 可以等同于一个元素 $T_v \in H^{-m}(\Omega)$. 以这种方式, 利用从空间 $L^2(\Omega)$ 到 $H^{-m}(\Omega)$ 中的典则内射, 空间 $L^2(\Omega)$ 就嵌入空间 $H^{-m}(\Omega)$.

在本节的其余部分, 我们将把注意力集中在空间

$$H^{-1}(\Omega) := (H_0^1(\Omega))'$$

上, 其中开集 Ω 是 $\mathbb{R}^N$ 中的一个区域, 这是因为在本章的各处, 该空间都将起着关键作用. 第一个性质是紧性. 就其精神实质而言, 是与 Rellich-Kondrachov 定理 (定理 6.6-3) 类似的, 故仍冠以此名是合理的.

定理 6.11-3 (在 $L^2(\Omega)$ 中的 Rellich-Kondrachov 紧嵌入定理) 设 Ω 是 $\mathbb{R}^N$ 中的区域. 则从 $L^2(\Omega)$ 到 $H^{-1}(\Omega)$ 中的典则内射是紧的.

证明 典则内射

$$v \in L^2(\Omega) \to T_v \in H^{-1}(\Omega)$$

只是典则内射

$$\iota : H_0^1(\Omega) \to L^2(\Omega)$$

的对偶算子 ι' (5.11 节). 在此, 空间 $L^2(\Omega)$ 等同于其对偶空间. 为说明这一点, 只需验证

$$T_v(w) = (v, \iota w), \quad 对所有\ v \in L^2(\Omega)\ 及所有\ w \in H_0^1(\Omega),$$

其中 $(\cdot, \cdot)$ 表示 $L^2(\Omega)$ 的内积; 等价地, 只需验证

$$T_v(w) = \int_\Omega vw dx, \quad 对所有\ v \in L^2(\Omega)\ 及所有\ w \in H_0^1(\Omega).$$

但这个表达式由 T_v 的定义及 $\mathcal{D}(\Omega)$ 在 $H_0^1(\Omega)$ 中的稠密性立即可以得到.

因为根据 Rellich-Kondrachov 嵌入定理 (定理 6.6-3), ι 是紧的, 由定理 5.11-2 知 ι' 也是紧的. □

设 Ω 为 $\mathbb{R}^N$ 的开子集. 由于一个函数 $v \in L^2(\Omega)$ 可等同其定义的分布 (如上所述), 显然

$$v \in L^2(\Omega)\ 意味着\ v \in H^{-1}(\Omega)\ 且\ \partial_i v \in H^{-1}(\Omega), \quad 1 \leqslant i \leqslant N,$$

这是因为

$$|T_v(\varphi)| = \left|\int_\Omega v\varphi dx\right| \leqslant \|v\|_{0,\Omega}\|\varphi\|_{1,\Omega}, \quad 对所有\ \varphi \in \mathcal{D}(\Omega),$$

$$|\partial_i T_v(\varphi)| = |-T_v(\partial_i\varphi)| = \left|-\int_\Omega v\partial_i\varphi dx\right|$$
$$\leqslant \|v\|_{0,\Omega}\|\varphi\|_{1,\Omega} \quad 对所有\ \varphi \in \mathcal{D}(\Omega).$$

值得关注, 但其证明又出乎意外的困难的是, 下述逆反结论成立 (这里要注意, 假设 $v \in \mathcal{D}'(\Omega)$ 较之 $v \in H^{-1}(\Omega)$ 更弱).

定理 6.11-4 (J. L. Lions 引理[47) 48)]) 设 Ω 是 $\mathbb{R}^N$ 中的区域. 则

$$v \in \mathcal{D}'(\Omega) \text{ 及 } \partial_i v \in H^{-1}(\Omega), \quad 1 \leqslant i \leqslant N, \text{ 意味着 } v \in L^2(\Omega).$$

J. L. Lions 引理具有极其重要的意义: 在本章的其余部分可以看到, 其在许多基本结果的证明中起着关键作用, 如 Stokes 方程的弱形式解的存在性 (6.14 节), Korn 不等式 (6.15 节), 弱 Poincaré 引理 (6.17 节), 弱 Saint-Venant 引理 (6.18 节) 以及弱 Donati 引理 (6.19 节) 等.

要注意, 在我们对 J. L. Lions 引理的所有应用中, 使得 $\partial_i v \in H^{-1}(\Omega), 1 \leqslant i \leqslant N$, 的分布 $v \in \mathcal{D}'(\Omega)$ 属于 $\mathcal{D}'(\Omega)$ 的一个严格的子空间, 如 $H^{-1}(\Omega)$ 或 $L^1_{\mathrm{loc}}(\Omega)$.

注 虽然在本书中将定理 6.11-4 称为 J. L. Lions 引理, 在文献中还有他的另一些结果被冠以同样的名称, 如 "紧性引理"[49)] 及 "奇摄动引理"[50)] 等.

最后, 我们介绍 J. L. Lions 引理的一个有用的推广.[51)]

定理 6.11-5 ($H^m(\Omega)$ 中的 J. L. Lions 引理) 设 Ω 是 $\mathbb{R}^N$ 中的区域, 而 $m \in \mathbb{Z}$ 是任意整数. 则

---

47) $v \in H^{-1}(\Omega)$ 及 $\partial_i v \in H^{-1}(\Omega), 1 \leqslant i \leqslant N$, 意味着 $v \in L^2(\Omega)$ 这一结论, 对于具有光滑边界的区域, 首先由 Jacques-Louis Lions (1928—2001) 确立, 可见下文脚注 (22):

E. Magenes; G. Stampacchia [1958]: I problemi al contorno per le equazioni differenziali di tipo ellitico. Annalidella Scuola Normale Superiore Pisa **12**, 247–358.

J. L. Lions 的证明最先发表于 Duvaut & Lions [1976] 中. 自此以后, 出现了关于这一结论的另一些证明, 其中一些将其推广到真区域 (即如定理 6.11-4 中所述, 具有 Lipschitz 连续边界的区域) 的情形, 另一些将其推广到把空间 $H^{-1}(\Omega)$ 换为更一般的空间 $W^{-1,q}(\Omega), 1 < q < \infty$, 的情形. 参见:

L. Tartar [1978]: Topics in Nonlinear Analysis. Publications Mathématiques d'Orsay No.78.13, Université de Paris-Sud, Orsay.

G. Geymonat; P. Suquet [1986]: Functional Spaces for Norton-Hoff materials. Mthematical Methods in the Applied Sciences 8, 206–222.

在下文中, 给出了一个当 Ω 不是区域时, J. L. Lions 引理的反例:

G. Geymonat; G. Gilardi [1998]: Contre-exemple à l'inégalité de Korn et an lemme de Lions dans des domaines irréguliers. Equations aux Dérivées et Appliations. Articles Dédiés à Jacques-Louis Lions, pp. 541–548, Gauthier-Villars, Paris.

48) 假设 $v \in H^{-1}(\Omega)$ 可以换为较弱的假设 $v \in \mathcal{D}'(\Omega)$ 这一结论由下文确立:

W. Borchers; H. Sohr [1990]: On the equations rot $v = g$ and div $u = f$ with boundary conditions. Hokkaido Mathematical Journal 19, 67–87.

49) 见 Lions [1961, 第 X 章, 命题 4.1] 或 Lions [1969, 第 X 章, 5.2 节].

50) 见 Lions [1973, 第 X 章, 引理 5.1].

51) 这一结果属于:

C. Amrouche; V. Girault [1994]: Decomposition of vector spaces and application to the Stokes problem in arbitrary dimension. Czechoslovak Mathematical Journal **44**, 109–140.

$$v \in \mathcal{D}'(\Omega) \text{ 且 } \partial_i v \in H^m(\Omega), \quad 1 \leqslant i \leqslant N,$$

意味着 $v \in H^{m+1}(\Omega)$.

习题

6.11-1 在定理 6.11-1 的假设下, 仍用其中的符号, 证明

$$\begin{aligned}\|\ell\|_{(W^{m,p}(\Omega))'} = \inf\Bigg\{ & \|(f^{\boldsymbol{\alpha}})\|_{(L^q(\Omega))^M} := \left(\sum_{|\boldsymbol{\alpha}|\leqslant m} \|f^{\boldsymbol{\alpha}}\|^q_{L^q(\Omega)}\right)^{1/q}; \\ & (f^{\boldsymbol{\alpha}}) \in (L^q(\Omega))^M \quad \text{对每个 } |\boldsymbol{\alpha}| \leqslant m \\ & \text{及} \sum_{|\boldsymbol{\alpha}|\leqslant m} \int_\Omega f_{\boldsymbol{\alpha}} \partial^{\boldsymbol{\alpha}} v dx = \ell(v) \quad \text{对所有 } v \in W^{m,p}(\Omega) \Bigg\},\end{aligned}$$

而且上面的下确界可以被积空间 $(L^q(\Omega))^M$ 中的一元素达到.

6.11-2 设 $1 < p < \infty$, 而 $(v_k)_{k=1}^\infty$ 是空间 $W^{m,p}(\Omega)$ 中的有界序列. 利用定理 6.11-1 及空间 $L^p(\Omega)$ 的自反性 (定理 5.14-2), 证明存在一个在空间 $W^{m,p}(\Omega)$ 中弱收敛的子序列 $(v_{\sigma(k)})_{k=1}^\infty$.

这一性质与 Banach-Eberlein-Šmulian 定理 (定理 5.14-4) 的 (b) 部分结合, 就提供了空间 $W^{m,p}(\Omega), 1 < p < \infty$, 自反性的另一个 (与定理 6.5-1 不同的) 证明.

6.12 Babuška-Brezzi 上下确界定理; 对有约束二次极小化问题的应用

J. L. Lions 引理的第一个应用是确立 Stokes 方程弱解的存在性 (见定理 6.14-1 的证明). 为此目的, 我们首先要建立一个抽象的泛函框架, 以此作为产生于流体及固体力学中的各种基本线性问题的模型. 这是本节的课题.

迄今为止, 我们讨论的是形如 "求 $u \in V$ 使得 $a(u,v) = \ell(v)$ 对所有 $v \in V$" (定理 6.2-1) 的线性抽象变分问题. 现在考虑另一类线性抽象变分问题, 其定义需要第二个空间及第二个双线性形式 (在以下定理中分别以 M 及 b 表示).

首先, 我们要确立这种问题基本的存在性及唯一性结果. 在下述定理中关于双线性形式 b 所作的假设, 即存在一个常数 β 使

$$\beta > 0 \quad \text{且} \quad \inf_{\left\{\substack{\mu \in M \\ \mu \neq 0}\right.} \sup_{\left\{\substack{v \in V \\ v \neq 0}\right.} \frac{|b(v,\mu)|}{\|v\|_V \|\mu\|_M} \geqslant \beta,$$

称为 Babuška-Brezzi 上下确界条件[52).]

注 Lax-Milgram 定理是下述定理的一个特殊情况 ($M=\{0\}, b=0$).

定理 6.12-1 (Babuška-Brezzi 上下确界定理) 设 V 和 M 是两个 Hilbert 空间, $a(\cdot,\cdot): V\times V\to\mathbb{R}$ 及 $b: V\times M\to\mathbb{R}$ 是具有下述性质的两个连续的双线性形式: 存在常数 α 使得

$$\alpha>0 \text{ 且 } a(v,v)\geqslant\alpha\|v\|_V^2 \quad \text{对所有 } v\in U_0:=\{v\in V; b(v,\mu)=0 \text{ 对所有 } \mu\in M\},$$

即 $a(\cdot,\cdot)$ 是 U_0 强制的, 另外还存在常数 β 使得

$$\beta>0 \text{ 且 } \inf_{\left\{\substack{\mu\in M\\ \mu\neq 0}\right.}\sup_{\left\{\substack{v\in V\\ v\neq 0}\right.}\frac{|b(v,\mu)|}{\|v\|_V\|\mu\|_M}\geqslant\beta.$$

最后, 还假设 $\ell: V\to\mathbb{R}$ 及 $\chi: M\to\mathbb{R}$ 是两个连续的线性形式.

则变分问题: 求 $(u,\lambda)\in V\times M$ 使得

$$\begin{aligned} a(u,v)+b(v,\lambda)&=\ell(v) \quad \text{对所有 } v\in V,\\ b(u,\mu)&=\chi(\mu) \quad \text{对所有 } \mu\in M, \end{aligned}$$

有且只有一个解, 而且以这种方式定义的线性算子 $(\ell,\chi)\in V'\times M'\to(u,\lambda)\in V\times M$ 是连续的.

[52)] 冠名源自:

I. Babuška [1971]: Error bound for finite element method. Numerische Mathematik **16**, 322–333.

F. Brezzi [1974]: On the existence, uniqueness and approximation of saddle point problems arising from Lagrange multipliers. Revue Française d'Automatique, Informatique, et Recherche Opérationnelle-Série Rouge **8**, 129–151.

在这两篇论文中, "Babuška-Brezzi" 条件实际上是用两种不同但等价的形式给出的 (定理 6.12-1 中的形式是 Brezzi [1974] 给出的). 在下述诸文中确立了两种形式的等价性. 例如,

L. Demkowicz [2000]: Babuška ⇔ Brezzi??. Technical Report, Texas Institute for Computational and Applied Mathematics, TICAM Seminar (October 31, 2000).

Franco Brezzi 还令人佩服地确立了 (仍是在 Brezzi [1974] 中) inf-sup 条件的必要性; 见习题 6.12-1.

实际上, 这个条件此前已经出现过, 虽然只是出现在一个特定问题的处理中 (即不是如定理 6.12-1 中那样的 "抽象" 形式), 见:

O. A. Ladyzhenskaya [1969]: The Mathematical Theory of Viscous Flows. Second Edition. Gordor and Breach, New york.

据此原因, 该条件有时称为 *Ladyzhenskaya-Babuška-Brezzi* 条件. 实际上, 形如 Babuška [1971] 中给出的结果, 在下文的定理 3.1 中已给出证明:

J. Nečas [1962]: Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle. Annali della Scuola Normale Superiore di Pisa, Classe di Scienze, Serie III, **16**, 305–326.

Ivo Babuška 和 Franco Brezzi 首先证明了, 这种类型的结果也是当对此类变分问题施行有限元逼近时, 基本误差估计的关键.

证明　对每一个 $u \in V$, 线性形式 $v \in V \to a(u,v) \in \mathbb{R}$ 是连续的. 所以, 存在唯一的元素 $Au \in V'$ 使得

$$a(u,v) = Au(v) \quad \text{对所有 } (u,v) \in V \times V.$$

此外,

$$\begin{aligned}\|Au\|_{V'} &= \sup_{v \neq 0} \frac{|Au(v)|}{\|v\|_V} = \sup_{v \neq 0} \frac{|a(u,v)|}{\|v\|_V} \\ &\leqslant \|a\|_{\mathcal{L}_2(V;\mathbb{R})}\|u\| \quad \text{对所有 } u \in V,\end{aligned}$$

故以这种方式定义, 显然线性的映射 $A : V \to V'$ 是连续的, 而且 $\|A\|_{\mathcal{L}(V;V')} \leqslant \|a\|_{\mathcal{L}_2(V;\mathbb{R})}$. 利用同样的推理可以证明, 存在一个映射 $B \in \mathcal{L}(V;M')$, 因而存在对偶算子 $B' \in \mathcal{L}(M;V')$ (5.11 节), 使得

$$b(v,\mu) = Bv(\mu) = B'\mu(v) \quad \text{对所有 } (v,\mu) \in V \times M.$$

这样, 求解定理 6.12-1 中的抽象变分问题就归结为寻求偶对 $(u,\lambda) \in V \times M$ 使其满足下述算子方程组:

$$\begin{aligned}Au + B'\lambda = \ell \quad &\text{在 } V' \text{ 中}, \\ Bu = \chi \quad &\text{在 } M' \text{ 中}.\end{aligned}$$

从对偶算子 B' 的角度来看, Babuška-Brezzi 上下确界条件等价于

$$\begin{aligned}\|B'\mu\|_{V'} &= \sup_{v \neq 0} \frac{|B'\mu(v)|}{\|v\|_V} \\ &= \sup_{v \neq 0} \frac{|b(v,\mu)|}{\|v\|_V} \geqslant \beta\|\mu\|_M \quad \text{对所有 } \mu \in M,\end{aligned}$$

这一关系式正是出现在 Banach 闭值域定理 (第二部分; 见定理 5.11-6) 中的三个等价条件 (在此应用于算子 $B \in \mathcal{L}(V;M')$) 之一.

所以我们从这个定理可以推断得, 算子 $B : V \to M'$ 是满射, 算子 $B' : M \to V'$ 是单射, 而且空间 $\operatorname{Im} B'$ 在 V' 中是闭的.

因为 B 是满射且 $\chi \in M'$, 故存在 $u_0 \in V$ 使得

$$Bu_0 = \chi.$$

由于

$$\begin{aligned}U_0 &= \{v \in V; b(v,\mu) = 0 \text{ 对所有 } \mu \in M\} \\ &= \{v \in V; Bv(\mu) = 0 \text{ 对所有 } \mu \in M\} \\ &= \{v \in V; Bv = 0 \text{ 在 } M' \text{ 中}\} = \operatorname{Ker} B,\end{aligned}$$

双线性形式 $a(\cdot,\cdot)$ 实际上是 Ker B 强制的. 所以根据 Lax-Milgram 定理 (定理 6.2-1; 显然线性形式 $v \in V \to \ell(v) - a(u_0, v)$ 是连续的), 存在唯一的 $u_1 \in \text{Ker } B$ 使得

$$a(u_1, v) = \ell(v) - a(u_0, v) \quad \text{对所有 } v \in \text{Ker } B.$$

于是元素

$$u := (u_0 + u_1) \in V$$

就满足

$$\begin{aligned}(\sigma(Au - \ell), v)_V &= (Au - \ell)(v) \\ &= a(u, v) - \ell(v) = 0 \quad \text{对所有 } v \in \text{Ker } B,\end{aligned}$$

其中 $(\cdot,\cdot)_V$ 与 $\sigma : V' \to V$ 分别表示空间 V 的内积及 F. Riesz 等距. 换言之,

$$\sigma(Au - \ell) \in (\text{Ker } B)^{\perp},$$

其中 $(\text{Ker } B)^{\perp}$ 表示 Ker B 在 Hilbert 空间 $(V, (\cdot,\cdot)_V)$ 中的正交补.

设 $(\cdot,\cdot)_M$ 及 $\tau : M' \to M$ 分别表示空间 M 的内积及 F. Riesz 等距. 则

$$\begin{aligned}(\tau Bv, \mu)_M &= Bv(\mu) = B'\mu(v) \\ &= (\sigma B'\mu, v)_V \quad \text{对所有 } v \in V \text{ 及所有 } \mu \in M,\end{aligned}$$

这说明 $\sigma B'$ 是 σB 在 Hilbert 空间意义下的伴随算子, 如在定理 4.7-2(a) 中所定义的.

因此, 由上述同一定理的 (b) 部分, 有

$$\text{Ker } \tau B \oplus \overline{\text{Im } \sigma B'} = V.$$

但因 Im B' 在 V' 中是闭的, 故 Im $\sigma B'$ 在 V 中是闭的, 而 $\sigma : V' \to V$ 是一个等距; 所以

$$\text{Im } \sigma B' = (\text{Ker } \tau B)^{\perp} = (\text{Ker } B)^{\perp}.$$

因为 $\sigma(Au - \ell) \in (\text{Ker } B)^{\perp}$, 故存在 $\lambda \in M$ 使得 $-\sigma B'\lambda = \sigma(Au - \ell)$ 或等价地

$$Au + B'\lambda = \ell.$$

另一方面, 因 $u_1 \in \text{Ker } B$,

$$Bu = B(u_0 + u_1) = Bu_0 = \chi,$$

因此定理 6.12-1 中的变分问题解 $(u, \lambda) \in V \times M$ 的存在性得以确立.

因为变分问题是线性的, 确立其解的唯一性就归结为证明, 如果 $(u, \lambda) \in V \times M$ 满足

$$\begin{aligned}a(u, v) + b(v, \lambda) &= 0 \quad \text{对所有 } v \in V, \\ b(u, \mu) &= 0 \quad \text{对所有 } \mu \in M,\end{aligned}$$

则 $(u,\lambda)=(0,0)$. 在第一个方程中令 $v=u$, 因 $b(u,\lambda)=0$, 有

$$a(u,u)+b(u,\lambda)=a(u,u)=0.$$

由于第二个变分方程意味着 $u\in \mathrm{Ker}\,B$, 而由假设双线性形式 $a(\cdot,\cdot)$ 是 $\mathrm{Ker}\,B$ 强制的, 故 $u=0$. 这样, 第一个方程就化为 $b(v,\lambda)=0$ 对所有 $v\in V$ 或等价地

$$Bv(\lambda)=B'\lambda(v)=0 \quad 对所有\ v\in V,$$

这说明 $B'\lambda=0$. 但 B' 是单射, 故 $\lambda=0$.

映射

$$\mathcal{A}:(v,\mu)\in V\times M\to\mathcal{A}(v,\mu):=(Av+B'\mu,Bv)\in V'\times M'$$

显然是连续的, 而且是双射. 由 Banach 开映射定理 (定理 5.6-1) 立得其逆映射 $\mathcal{A}^{-1}:V'\times M'\to V\times M$ 的连续性. □

注 (1) Banach 闭值域定理证明, Babuška-Brezzi 上下确界条件当且仅当映射 $B\in\mathcal{L}(V;M')$ 是满射, 或者当且仅当映射 $B'\in\mathcal{L}(M;V')$ 是单射且在 V' 中有闭值域时成立.

(2) Babuška-Brezzi 上下确界条件对定理 6.12-1 中变分问题解的存在性也是必要的, 特别地, 存在性意味着对任意的 $\chi\in M'$, 方程 $Bu=\chi$ 必定有一个解 $u\in V$ (如在上述证明中所示), 即映射 $B\in\mathcal{L}(V;M')$ 必须是满射; 关于这方面, 也可见习题 6.12-1, 其中 inf-sup 条件的必要性通常可以确立.

在双线性形式 $a(\cdot,\cdot)$ 是对称的这一附加假设下, 我们下面将证明, 定理 6.12-1 提供了一个很有意义的方法, 用于求解在 6.1 节中曾讨论过的特定的二次极小问题. 此时, Hilbert 空间 V 的那个非空闭凸子集, 在其上极小化形如

$$J:v\in V\to J(v):=\frac{1}{2}a(v,v)-\ell(v)$$

的二次泛函, 将被表示为如下形式

$$U_\chi:=\{v\in V;b(v,\mu)=\chi(\mu)\ 对所有\ \mu\in M\},$$

其中 M 是另一个 Hilbert 空间, 而 $b:V\times M\to\mathbb{R}$ 与 $\chi:M\to\mathbb{R}$ 则是满足 Babuška-Brezzi 上下确界定理假设的连续双线性形式与线性形式.

这种极小化问题提供一个约束二次极小化问题的例子, 约束体现在, 任何极小化解 u (如果存在的话) 都应满足

$$b(u,\mu)=\chi(\mu)\quad 对所有\ \mu\in M.$$

定理 6.12-2 设关于空间 V 和 M 的假设, 关于双线性形式 $a(\cdot,\cdot):V\times V\to\mathbb{R}$ 和 $b(\cdot,\cdot):V\times M\to\mathbb{R}$ 的假设, 以及关于线性形式 $\ell:V\to\mathbb{R}$ 和 $\chi:M\to\mathbb{R}$ 的假设都与定理 6.12-1 中的相同. 另外还假定双线性形式 $a(\cdot,\cdot)$ 是对称的.

则 $(u, \lambda) \in V \times M$ 是定理 6.12-1 中变分问题的唯一解, 即

$$a(u, v) + b(v, \lambda) = \ell(v) \quad \text{对所有 } v \in V,$$
$$b(u, \mu) = \chi(\mu) \quad \text{对所有 } \mu \in M,$$

其充分必要条件为, u 是下述约束二次极小化问题的唯一解:

$$u \in U_\chi \quad \text{且} \quad J(u) = \inf_{v \in U_\chi} J(v),$$

其中空间 V 的子集 U_χ 及泛函 $J : V \to \mathbb{R}$ 分别由以下两式定义:

$$U_\chi := \{v \in V; b(v, \mu) = \chi(\mu) \text{ 对所有 } \mu \in M\},$$
$$J(v) := \frac{1}{2}a(v, v) - \ell(v) \text{ 对每一个 } v \in V.$$

证明 设 $(u, \lambda) \in V \times M$ 是定理 6.12-1 中变分问题的唯一解. 那么第二个变分方程说明 $u \in U_\chi$.

双线性形式 $a(\cdot, \cdot)$ 的对称性 (这一假设在此是实质性的) 意味着

$$J(u + w) - J(u) = (a(u, w) - \ell(w)) + \frac{1}{2}a(w, w) \quad \text{对所有 } w \in V.$$

但是第一个变分方程意味着

$$a(u, w) - \ell(w) = -b(w, \lambda) = 0 \quad \text{对所有 } w \in U_0 := \{v \in V; b(v, \mu) = 0 \text{ 对所有 } \mu \in M\}.$$

所以

$$J(u + w) - J(u) = \frac{1}{2}a(w, w) \geqslant \frac{\alpha}{2}\|w\|^2 > 0 \quad \text{对所有 } w \in U_0, w \neq 0$$

(因 $a(\cdot, \cdot)$ 是 U_0 强制的), 这就证明了 $J(u) = \inf_{v \in U_\chi} J(v)$ 而且 $u \in U_\chi$ 是这个约束二次极小化问题的唯一解.

反之, 假设 $\widetilde{u} \in U_\chi$ 满足 $J(\widetilde{u}) = \inf_{v \in U_\chi} J(v)$, 又设 $(u, \lambda) \in V \times M$ 是定理 6.12-1 中变分问题的唯一解. 则上述推理说明, $u \in U_\chi$ 且 $J(u) = \inf_{v \in U_\chi} J(v)$. 因为这个极小化问题的解是唯一的 (由前述可看出, 如果 $v \in U_\chi$ 且 $v \neq u$ 则 $J(v) > J(u)$), 所以 $\widetilde{u} = u$. □

利用定理 6.12-2, 可以将求解一特定的约束问题化为求解一非约束问题. 或许最有意义的是线性形式 χ 为零的这一特殊情况, 因为在这种情况下 (在 6.1 节中所讨论的) 下列约束变分问题: 求 $u \in U_0 = \{v \in V; b(v, \mu) = 0 \text{ 对所有 } \mu \in M\}$ 使得

$$a(u, v) = \ell(v) \quad \text{对所有 } v \in U_0.$$

(定理 6.1-1) 就转换为无约束的变分问题: 求 $(u, \lambda) \in V \times M$ 使得

$$a(u, v) + b(v, \lambda) = \ell(v) \quad \text{对所有 } v \in V,$$
$$b(u, \mu) = 0 \quad \text{对所有 } \mu \in M.$$

在此, "无约束" 反映在变分方程 $a(u,v)+b(v,\lambda)=\ell(v)$ 是对全空间 V 中所有的 v 都满足, 而方程 $a(u,v)=\ell(v)$ 只是对 V 的子空间 U_0 中所有 v 满足, U_0 是由约束 $b(v,\mu)=0$ 对所有 $\mu\in M$ 来定义的.

定理 6.12-1 及定理 6.12-2 的第一个应用 (用于 $\mathbb{R}^n$ 中的一个约束二次极小化问题) 在习题 6.12-2 中给出. 其他的应用可在以下两节中找到.

注 我们将证明, 在以下附加假设下: $a(v,v)\geqslant 0$ 对所有 $v\in V$, 在定理 6.12-2 中的偶对 (u,λ) 是一个特定的 Lagrange 泛函 $\mathcal{L}:V\times M\to\mathbb{R}$ 的鞍点, 而第二个变量 $\lambda\in M$ 则是与约束 $b(u,\mu)=\chi(\mu)$ (对所有 $\mu\in M$) 相应的 Lagrange 乘子 (7.16 节).

习题

6.12-1 设 V 和 M 是两个 Hilbert 空间, 而 $a(\cdot,\cdot):V\times V\to\mathbb{R}$ 和 $b(\cdot,\cdot):V\times M\to\mathbb{R}$ 是两个连续的双线性形式.

(1) 假定任意给出两个连续线性形式 $\ell:V\to\mathbb{R}$ 和 $\chi:M\to\mathbb{R}$ 都存在一个且只有一个偶对 $(u,\lambda)\in V\times M$ 满足变分方程

$$a(u,v)+b(v,\lambda)=\ell(v)\quad \text{对所有 } v\in V$$

和

$$b(u,\mu)=\chi(\mu)\quad \text{对所有 } \mu\in M.$$

证明定理 6.12-1 证明中定义的算子 $A\in\mathcal{L}(V;V')$ 和 $B\in\mathcal{L}(V;M')$ 必然具有下述两条性质: A 是从 V 到 V' 上的双射而 $\mathrm{Im}\,B=M'$.

首先, 设算子 $\rho\in\mathcal{L}\,(V',(\mathrm{Ker}\,B)')$ 对所有 $v'\in V'$ 及所有 $v\in\mathrm{Ker}\,B$, 由 $(\rho v')(v)=v'(v)$ 及所有 $v\in\mathrm{Ker}\,B$ 定义. 那么算子 $\rho A\in\mathcal{L}(V;(\mathrm{Ker}\,B)')$ 到 $\mathrm{Ker}\,B$ 的限制是具有连续逆的双射 (在定理 6.12-1 中所作的假设, 即双线性形式 $a(\cdot,\cdot)$ 是 $\mathrm{Ker}\,B$ 强制的, 显然意味着这一性质成立).

其次, inf-sup 条件满足 (如定理 6.12-1 中的证明所示, 这实际上是关于算子 B 的对偶算子 B' 的一个假定).

(2) 反之, 假定算子 $A\in\mathcal{L}(V;V')$ 及 $B\in\mathcal{L}(V;M')$ 使得上述两条性质满足. 证明, 给定任意两个连续线性形式 $\ell:V\to\mathbb{R}$ 及 $\chi:M\to\mathbb{R}$, 问题 (1) 的变分方程组存在一个且只有一个解 $(u,\lambda)\in V\times M$[53].

注 问题 (2) 包含定理 6.12-2 作为一个特殊情况.

6.12-2 设 $\boldsymbol{A}$ 是 $n\times n$ 实对称矩阵, 而 $\boldsymbol{B}$ 是秩为 m 的实 $m\times n$ 矩阵 (因此 $m\leqslant n$) 且具有以下性质: 存在 $\alpha>0$ 使得 $\boldsymbol{v}^T\boldsymbol{A}\boldsymbol{v}\geqslant\alpha\boldsymbol{v}^T\boldsymbol{v}$ 对所有 $\boldsymbol{v}\in\mathrm{Ker}\,\boldsymbol{B}$ (因此是一个较 $\boldsymbol{A}$ 的正定性弱的性质).

---

[53] 这个结果归于 Brezzi [1974].

证明, 给定任意的 $\boldsymbol{c} \in \mathbb{R}^n$ 和 $\boldsymbol{d} \in \mathbb{R}^m$, 线性方程组 (其矩阵是 $n+m$ 阶对称矩阵)

$$\begin{aligned} \boldsymbol{Au} + \boldsymbol{B}^T\boldsymbol{\lambda} &= \boldsymbol{c}, \\ \boldsymbol{Bu} &= \boldsymbol{d}, \end{aligned}$$

有唯一解 $(\boldsymbol{u}, \boldsymbol{\lambda}) \in \mathbb{R}^n \times \mathbb{R}^m$, 而且 $\boldsymbol{u}$ 是下述 $\mathbb{R}^n$ 中约束二次极小化问题的唯一解: 求

$$\boldsymbol{u} \in \boldsymbol{U_d} = \{\boldsymbol{v} \in \mathbb{R}^n; \boldsymbol{Bv} = \boldsymbol{d}\}$$

使得

$$J(\boldsymbol{u}) = \inf_{\boldsymbol{v} \in \boldsymbol{U_d}} J(\boldsymbol{v}),$$

其中

$$J(\boldsymbol{v}) := \frac{1}{2}\boldsymbol{v}^T\boldsymbol{Av} - \boldsymbol{c}^T\boldsymbol{v} \quad \text{对所有 } \boldsymbol{v} \in \mathbb{R}^n.$$

注 如果 $\boldsymbol{B} = \boldsymbol{0}$ 及 $\boldsymbol{d} = \boldsymbol{0}$, 在这种情况下 $\boldsymbol{A}$ 是正定的且 $\boldsymbol{U_d} = \mathbb{R}^n$, 上述极小化问题的解 $\boldsymbol{u} \in \mathbb{R}^n$ 也是 n 阶线性方程组 $\boldsymbol{Au} = \boldsymbol{c}$ 的解. 值得注意的是, 在这里所讨论的更一般情况下, $\boldsymbol{u}$ 仍旧能由求解*线性方程组*得到, 但此时是 $n+m$ 阶方程组. 稍后 (7.15 节) 我们将会看到, 出现在这个线性方程组中的*辅助未知量* $\boldsymbol{\lambda} \in \mathbb{R}^m$ 实际上是与*约束* $\boldsymbol{Bv} = \boldsymbol{d}$ 相应的 Lagrange *乘子*.

6.13 Babuška-Brezzi 上下确界定理的应用: 变分问题的原始, 混合及对偶形式

要注意将本节中所定义的 "原始形式" 及 "对偶形式" 与将在 7.16 节中定义的 "原始问题" 及 "对偶问题" 区别开来.

在这一节中, 我们用下述的*模型问题*来说明定理 6.12-1 和定理 6.12-2 的应用. 该模型问题相应于 $-\Delta$ 的齐次 Dirichlet 问题 (6.7 节): 求 $u \in H_0^1(\Omega)$ 使得

$$\int_\Omega \sum_{i=1}^n \partial_i u \partial_i v dx = \int_\Omega f v dx \quad \text{对所有 } v \in H_0^1(\Omega),$$

其中 Ω 是 $\mathbb{R}^N$ 中的区域, 而 $f \in L^2(\Omega)$ 是给定的函数. 前面已证明 (定理 6.7-2), 这个问题有唯一解, 它也是下述二次极小化问题的唯一解: 求 $u \in H_0^1(\Omega)$ 使得

$$J(u) = \inf_{v \in H_0^1(\Omega)} J(v),$$

其中

$$J(v) := \frac{1}{2}\int_\Omega \sum_{i=1}^N |\partial_i v|^2 dx - \int_\Omega f v dx.$$

在本节中, 这个极小化问题被视为 (模型问题的) *原始形式*.

然而, 在有些应用中显示, 不是函数 u 本身, 而是向量场

$$\mathbf{grad}\, u := (\partial_i u)_{i=1}^N \in \boldsymbol{L}^2(\Omega) := L^2(\Omega; \mathbb{R}^N)$$

是更值得关注的未知量. 这样, 自然就产生一个问题, 向量场 $\mathbf{grad}\, u$ 是否能像函数 u 那样, 直接被看作一个特定极小化问题的解? 如下面的定理中所阐明的, 要肯定地回答这个问题, 包含两个步骤:

第一, 构造一个形如定理 6.12-1 中所讨论的变分问题, 同时以 u 及 $\mathbf{grad}\, u$ 作为未知量; 这个问题构成 (模型问题) 的混合变分形式.

第二, 定理 6.12-2 给出一个只以 $\mathbf{grad}\, u$ 作为未知量的, 约束二次极小化问题; 它构成 (模型问题的) 对偶形式.

在下面, $\boldsymbol{a} \cdot \boldsymbol{b}$ 表示两个向量 $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^N$ 的 Euclid 内积, $|\boldsymbol{a}| := \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}}$ 表示向量 $\boldsymbol{a} \in \mathbb{R}^N$ 的 Euclid 范数, 而 $\|\cdot\|_{0,\Omega}$ 表示由下式定义的空间 $\boldsymbol{L}^2(\Omega)$ 中的积范数:

$$\|\boldsymbol{p}\|_{0,\Omega} := \left(\sum_{i=1}^N \|p_i\|_{0,\Omega}^2\right)^{1/2} \quad \text{对每个 } \boldsymbol{p} = (p_i)_{i=1}^N \in \boldsymbol{L}^2(\Omega).$$

下面两个定理对于同一个模型问题提供了两种不同的混合及对偶形式 (见定理 6.13-1 和定理 6.13-2 中的 (b) 和 (c) 部分, 为了方便起见模型问题的原始形式在以上两定理的 (a) 部分重述).

定理 6.13-1 ($-\Delta$ 的齐次 Dirchlet 问题的混合及对偶形式, 第一种情况) 设 Ω 是 $\mathbb{R}^N$ 中的区域, 函数 $f \in L^2(\Omega)$ 是给定的. 则

(a) 下述二次极小化问题存在唯一解 $u \in H_0^1(\Omega)$:

$$J(u) = \inf_{v \in H_0^1(\Omega)} J(v),$$

其中

$$J(v) := \frac{1}{2}\int_\Omega |\mathbf{grad}\, v|^2 dx - \int_\Omega fv dx \quad \text{对所有 } v \in H_0^1(\Omega).$$

(b) 存在唯一的偶对 $(\boldsymbol{p}, \lambda) \in \boldsymbol{L}^2(\Omega) \times H_0^1(\Omega)$ 满足变分问题:

$$\int_\Omega \boldsymbol{p} \cdot \boldsymbol{q} dx - \int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \lambda dx = 0 \quad \text{对所有 } \boldsymbol{q} \in \boldsymbol{L}^2(\Omega),$$
$$\int_\Omega \boldsymbol{p} \cdot \mathbf{grad}\, \mu dx = \int_\Omega f\mu dx \quad \text{对所有 } \mu \in H_0^1(\Omega).$$

另外

$$\boldsymbol{p} = \mathbf{grad}\, u, \quad \lambda = u,$$

其中 $u \in H_0^1(\Omega)$ 是 (a) 中极小化问题的解.

(c) 向量场 $\boldsymbol{p} = \mathbf{grad}\, u$ 是下述约束二次极小化问题的唯一解:

$$\boldsymbol{p} \in \boldsymbol{U}_f := \{\boldsymbol{q} \in \boldsymbol{L}^2(\Omega); \int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \mu dx = \int_\Omega f \mu dx \quad \text{对所有 } \mu \in H_0^1(\Omega)\},$$

$$I(\boldsymbol{p}) = \inf_{\boldsymbol{q} \in \boldsymbol{U}_f} I(\boldsymbol{q}), \text{ 其中 } I(\boldsymbol{q}) := \frac{1}{2} \int_\Omega |\boldsymbol{q}|^2 dx \quad \text{对所有 } \boldsymbol{q} \in \boldsymbol{L}^2(\Omega).$$

证明 设双线性形式 $a(\cdot,\cdot) : \boldsymbol{L}^2(\Omega) \times \boldsymbol{L}^2(\Omega) \to \mathbb{R}, b(\cdot,\cdot) : \boldsymbol{L}^2(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ 及线性形式 $\ell : \boldsymbol{L}^2(\Omega) \to \mathbb{R}, \chi : H_0^1(\Omega) \to \mathbb{R}$, 分别由以下诸式定义:

$$a(\boldsymbol{p}, \boldsymbol{q}) := \int_\Omega \boldsymbol{p} \cdot \boldsymbol{q} dx \quad \text{对每个 } \boldsymbol{p}, \boldsymbol{q} \in \boldsymbol{L}^2(\Omega),$$

$$b(\boldsymbol{q}, \mu) := -\int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \mu dx \quad \text{对每个 } (\boldsymbol{q}, \mu) \in \boldsymbol{L}^2(\Omega) \times H_0^1(\Omega),$$

$$\ell := 0 \quad \text{而 } \chi(\mu) := \int_\Omega f \mu dx \quad \text{对每个 } \mu \in H_0^1(\Omega),$$

又设空间 $H_0^1(\Omega)$ 装备范数 $|\cdot|_{1,\Omega}$ (定理 6.5-2).

对于任意 $\mu \in H_0^1(\Omega)$, 向量场 $\boldsymbol{q}_\mu := \mathbf{grad}\, \mu$ 属于空间 $\boldsymbol{L}^2(\Omega)$ 且 $\|\boldsymbol{q}_\mu\|_{0,\Omega} = |\mu|_{1,\Omega}$. 这样, 对每个非零 $\mu \in H_0^1(\Omega)$ 就有

$$\sup_{\left\{\begin{smallmatrix} \boldsymbol{q} \in \boldsymbol{L}^2(\Omega) \\ \boldsymbol{q} \neq 0 \end{smallmatrix}\right.} \frac{\int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \mu dx}{\|\boldsymbol{q}\|_{0,\Omega}} \geqslant \frac{\left|\int_\Omega \boldsymbol{q}_\mu \cdot \mathbf{grad}\, \mu dx\right|}{\|\boldsymbol{q}_\mu\|_{0,\Omega}} = |\mu|_{1,\Omega},$$

这说明定理 6.12-1 中的 Babuška-Brezzi 上下确界条件成立, 在此

$$V := \boldsymbol{L}^2(\Omega), \quad M := H_0^1(\Omega).$$

定理 6.12-1 中的所有假设显然都成立. 所以 (b) 中的变分问题有唯一解 $(\boldsymbol{p}, \lambda) \in \boldsymbol{L}^2(\Omega) \times H_0^1(\Omega)$.

$\boldsymbol{p} = \mathbf{grad}\, u$ 和 $\lambda = \mu$ 显然满足 (b) 中变分问题的第一个方程. 同一问题中的第二个方程也同样由 $\boldsymbol{p} = \mathbf{grad}\, u$ 所满足, 这是因为 (a) 中的极小化问题的唯一解 $u \in H_0^1(\Omega)$ 也是以下变分方程的解:

$$\int_\Omega \mathbf{grad}\, u \cdot \mathbf{grad}\, \mu dx = \int_\Omega f \mu dx \quad \text{对所有 } \mu \in H_0^1(\Omega).$$

所以 (b) 得证.

最后, 由定理 6.12-2 立得 (c). 该定理在此适用是因为双线性形式 $a(\cdot,\cdot)$ 是对称的. □

注 因为空间 $\mathcal{D}(\Omega)$ 在空间 $H_0^1(\Omega)$ 中是稠密的, 而且对每个 $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$, 线性形式 $\mu \in H_0^1(\Omega) \to \int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \mu dx - \int_\Omega f \mu dx$ 是连续的, 故出现在定理 6.13-1(c) 中的

集合 $\boldsymbol{U}_f$ 实际上是由在分布意义下 (6.3 节) 满足下述偏微分方程的所有 $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$ 组成的:

$$\operatorname{div} \boldsymbol{q} + f = 0 \quad \text{在 } \mathcal{D}'(\Omega) \text{ 中}.$$

为了下一定理的需要, 我们首先定义一个向量场的空间: 给定 $\mathbb{R}^N$ 中任一开子集 Ω, 设

$$\boldsymbol{H}(\text{div}; \Omega) := \{\boldsymbol{q} \in \boldsymbol{L}^2(\Omega); \operatorname{div} \boldsymbol{q} \in L^2(\Omega)\},$$

其中 $\operatorname{div} \boldsymbol{q} := \sum_{i=1}^N \partial_i q_i$ 对每个 $\boldsymbol{q} = (q_i)_{i=1}^N \in \boldsymbol{L}^2(\Omega) = L^2(\Omega; \mathbb{R}^N)$. 就像我们在 Sobolev 空间 $H^1(\Omega)$ 的定义 (6.5 节) 中所见到的关系式 $\partial_i v \in \boldsymbol{L}^2(\Omega), 1 \leqslant i \leqslant N$, 关系式 “$\operatorname{div} \boldsymbol{q} \in L^2(\Omega)$” 也可理解为在分布的意义下成立. 这意味着一个向量场 $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$ 属于空间 $\boldsymbol{H}(\text{div}; \Omega)$ 当且仅当在 $\boldsymbol{L}^2(\Omega)$ 中存在 (唯一的) 函数, 以 $\operatorname{div} \boldsymbol{q}$ 表示, 满足

$$\int_\Omega (\operatorname{div} \boldsymbol{q}) \varphi dx = -\int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \varphi dx \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega).$$

容易验证 (利用类似于定理 6.5-1 中相关问题的证明方法), 装备如下定义的范数 $\|\cdot\|_{\boldsymbol{H}(\text{div};\Omega)}$:

$$\|\boldsymbol{q}\|_{\boldsymbol{H}(\text{div};\Omega)} := (\|\boldsymbol{q}\|_{0,\Omega}^2 + \|\operatorname{div} \boldsymbol{q}\|_{0,\Omega}^2)^{1/2} \quad \text{对每个 } \boldsymbol{q} \in \boldsymbol{H}(\text{div}; \Omega),$$

空间 $\boldsymbol{H}(\text{div}; \Omega)$ 是 Hilbert 空间[54].

要注意, 在定理 6.13-1 的证明中, 以 $a(\cdot, \cdot)$ 表示的双线性形式在全空间 $\boldsymbol{L}^2(\Omega)$ 上显然是强制的, 而不是仅在 $\boldsymbol{L}^2(\Omega)$ 的子空间 $\left\{\boldsymbol{q} \in \boldsymbol{L}^2(\Omega); \int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \mu dx = 0 \text{ 对所有 } \mu \in H_0^1(\Omega)\right\}$ 上强制. 在下一个证明中, 以 $a(\cdot, \cdot)$ 表示的双线性形式仅在空间 $\boldsymbol{H}(\text{div}; \Omega)$ 的真子空间 $\{\boldsymbol{q} \in \boldsymbol{H}(\text{div}; \Omega); \operatorname{div} \boldsymbol{q} = 0 \text{ 在 } \Omega \text{ 中 }\}$ 上强制.

定理 6.13-2 ($-\Delta$ 的齐次 Dirichlet 问题的混合及对偶形式, 第二种情况) 设 Ω 为 $\mathbb{R}^N$ 中的区域, 函数 $f \in \boldsymbol{L}^2(\Omega)$ 是给定的. 则

(a) 下述二次极小化问题

$$J(u) = \inf_{v \in H_0^1(\Omega)} J(v)$$

存在唯一解 $u \in H_0^1(\Omega)$, 其中

$$J(v) := \frac{1}{2} \int_\Omega |\mathbf{grad}\, v|^2 dx - \int_\Omega f v dx \quad \text{对每个 } v \in H_0^1(\Omega).$$

---

[54] 空间 $\boldsymbol{H}(\text{div}; \Omega)$ 以及其他有关的空间具有重要意义, 这些概念很自然地产出于各种具有重要物理意义问题的数学模型中. 空间 $\boldsymbol{H}(\text{div}; \Omega)$ 更深入的性质, 如特定的 Green 公式, 光滑函数的稠密性等, 参阅 Girault & Raviart [1986, 第 1 章].

(b) 存在唯一的偶对 $(\boldsymbol{p},\lambda)\in\boldsymbol{H}(\mathrm{div};\Omega)\times\boldsymbol{L}^2(\Omega)$ 满足变分问题:

$$\int_\Omega \boldsymbol{p}\cdot\boldsymbol{q}dx+\int_\Omega(\operatorname{div}\boldsymbol{q})\lambda dx=0 \quad \text{对所有 } \boldsymbol{q}\in\boldsymbol{H}(\mathrm{div};\Omega),$$
$$\int_\Omega(\operatorname{div}\boldsymbol{p})\mu dx=-\int_\Omega f\mu dx \quad \text{对所有 } \mu\in L^2(\Omega).$$

此外,

$$\boldsymbol{p}=\mathbf{grad}\, u,\quad \lambda=u,$$

其中 $u\in H_0^1(\Omega)$ 是 (a) 中极小化问题的解.

(c) 向量场 $\boldsymbol{p}=\mathbf{grad}\, u$ 是下述约束二次极小化问题的唯一解:

$$\boldsymbol{p}\in\widetilde{\boldsymbol{U}}_f:=\{\boldsymbol{q}\in\boldsymbol{H}(\mathrm{div};\Omega);\operatorname{div}\boldsymbol{q}+f=0 \quad \text{在 } L^2(\Omega)\text{中}\},$$
$$I(\boldsymbol{p})=\inf_{\boldsymbol{q}\in\widetilde{\boldsymbol{U}}_f} I(\boldsymbol{q}),\ \text{其中 } I(\boldsymbol{q}):=\frac{1}{2}\int_\Omega|\boldsymbol{q}|^2dx \quad \text{对每个 } \boldsymbol{q}\in\boldsymbol{L}^2(\Omega).$$

证明 设双线性形式 $a(\cdot,\cdot):\boldsymbol{H}(\mathrm{div};\Omega)\times\boldsymbol{H}(\mathrm{div};\Omega)\to\mathbb{R},\widetilde{b}(\cdot,\cdot):\boldsymbol{H}(\mathrm{div};\Omega)\times\boldsymbol{L}^2(\Omega)\to\mathbb{R}$, 及线性形式 $\ell:\boldsymbol{H}(\mathrm{div};\Omega)\to\mathbb{R},\chi:L^2(\Omega)\to\mathbb{R}$ 分别由以下诸式定义:

$$a(\boldsymbol{p},\boldsymbol{q}):=\int_\Omega\boldsymbol{p}\cdot\boldsymbol{q}dx \quad \text{对每个 } \boldsymbol{p},\boldsymbol{q}\in\boldsymbol{H}(\mathrm{div};\Omega),$$
$$\widetilde{b}(\boldsymbol{q},\mu):=\int_\Omega(\operatorname{div}\boldsymbol{q})\mu dx \quad \text{对每对 } (\boldsymbol{q},\mu)\in\boldsymbol{H}(\mathrm{div};\Omega)\times\boldsymbol{L}^2(\Omega),$$
$$\ell:=0 \text{ 而 } \chi(\mu):=\int_\Omega f\mu dx \quad \text{对每个 } \mu\in\boldsymbol{L}^2(\Omega).$$

双线性形式 $a(\cdot,\cdot)$ 在 $\boldsymbol{H}(\mathrm{div};\Omega)$ 的下述子空间上是强制的:

$$\widetilde{\boldsymbol{U}}_0=\left\{\boldsymbol{q}\in\boldsymbol{H}(\mathrm{div};\Omega);\int_\Omega(\operatorname{div}\boldsymbol{q})\mu dx=0 \text{ 对所有 } \mu\in\boldsymbol{L}^2(\Omega)\right\}$$
$$=\{\boldsymbol{q}\in\boldsymbol{H}(\mathrm{div};\Omega);\operatorname{div}\boldsymbol{q}=0 \text{ 在 } \boldsymbol{L}^2(\Omega) \text{ 中}\},$$

此因

$$a(\boldsymbol{q},\boldsymbol{q})=\|\boldsymbol{q}\|_{0,\Omega}^2=\|\boldsymbol{q}\|_{\boldsymbol{H}(\mathrm{div};\Omega)}^2 \quad \text{对所有 } \boldsymbol{q}\in\widetilde{\boldsymbol{U}}_0.$$

给定任意函数 $\mu\in\boldsymbol{L}^2(\Omega)$, 存在唯一的函数 $w_\mu\in H_0^1(\Omega)$ 满足

$$\int_\Omega\mathbf{grad}\, w_\mu\cdot\mathbf{grad}\, vdx=\int_\Omega\mu vdx \quad \text{对所有 } v\in H_0^1(\Omega).$$

特别地有 $\int_\Omega\mu\varphi dx=-\int_\Omega(-\mathbf{grad}\, w_\mu)\cdot\mathbf{grad}\,\varphi dx$ 对所有 $\varphi\in\mathcal{D}(\Omega)$. 根据空间 $\boldsymbol{H}(\mathrm{div};\Omega)$ 的定义, 这说明

$$\mathbf{grad}\, w_\mu\in\boldsymbol{H}(\mathrm{div};\Omega) \quad \text{且} \quad -\operatorname{div}\mathbf{grad}\, w_\mu=\mu\in L^2(\Omega).$$

因为存在常数 C 使得 $\|\mathbf{grad}\, w_\mu\|_{0,\Omega} \leqslant C\|\mu\|_{0,\Omega}$ 对所有 $\mu \in L^2(\Omega)$ (定理 6.7-2), 进一步得

$$\begin{aligned}\|\mathbf{grad}\, w_\mu\|_{\boldsymbol{H}(\mathrm{div};\Omega)} &= (\|\mathbf{grad}\, w_\mu\|_{0,\Omega}^2 + \|\mathrm{div}\,\mathbf{grad}\, w_\mu\|_{0,\Omega}^2)^{1/2} \\ &= (\|\mathbf{grad}\, w_\mu\|_{0,\Omega}^2 + \|\mu\|_{0,\Omega}^2)^{1/2} \\ &\leqslant \sqrt{C^2+1}\|\mu\|_{0,\Omega}.\end{aligned}$$

所以, 对每个非零 $\mu \in L^2(\Omega)$, 有

$$\begin{aligned}\sup_{\left\{\begin{subarray}{l}\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div};\Omega)\\ \boldsymbol{q} \neq \mathbf{0}\end{subarray}\right.} \frac{\left|\int_\Omega (\mathrm{div}\,\boldsymbol{q})\mu dx\right|}{\|\boldsymbol{q}\|_{\boldsymbol{H}(\mathrm{div};\Omega)}} &\geqslant \frac{\left|\int_\Omega (\mathrm{div}\,\mathbf{grad}\, w_\mu)\mu dx\right|}{\|\mathbf{grad}\, w_\mu\|_{\boldsymbol{H}(\mathrm{div};\Omega)}} \\ &\geqslant (C^2+1)^{-1/2}\|\mu\|_{0,\Omega}.\end{aligned}$$

这证明了定理 6.12-1 中的 Babuška-Brezzi 上下确界条件成立, 其中

$$V := \boldsymbol{H}(\mathrm{div};\Omega), \quad M := L^2(\Omega).$$

定理 6.12-1 中的所有其他假设显然都满足. 因此 (b) 中的变分问题有唯一解 $(\boldsymbol{p},\lambda) \in \boldsymbol{H}(\mathrm{div};\Omega) \times L^2(\Omega)$.

根据定义, 任一向量场 $\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div};\Omega)$ 满足

$$\int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \varphi dx + \int_\Omega (\mathrm{div}\,\boldsymbol{q})\varphi dx = 0 \quad \text{对所有 } \varphi \in \mathcal{D}(\Omega).$$

由于对每个 $\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div};\Omega)$, 线性形式 $\varphi \in \mathcal{D}(\Omega) \to \int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, \varphi dx + \int_\Omega (\mathrm{div}\,\boldsymbol{q})\varphi dx$ 是连续的, 而且 $\overline{\mathcal{D}(\Omega)} = H_0^1(\Omega)$, 就有

$$\int_\Omega \boldsymbol{q} \cdot \mathbf{grad}\, v dx + \int_\Omega (\mathrm{div}\,\boldsymbol{q})v dx = 0 \quad \text{对所有 } \boldsymbol{q} \in \boldsymbol{H}(\mathrm{div};\Omega) \text{ 及所有 } v \in H_0^1(\Omega).$$

特别地, 在上式中令 $v = u$, 就得到

$$\int_\Omega \mathbf{grad}\, u \cdot \boldsymbol{q} dx + \int_\Omega (\mathrm{div}\,\boldsymbol{q})u dx = 0 \quad \text{对所有 } \boldsymbol{q} \in \boldsymbol{H}(\mathrm{div};\Omega).$$

所以 $\boldsymbol{p} = \mathbf{grad}\, u$ 和 $\lambda = u$ 满足 (b) 中变分问题的第一个方程.

变分方程当 $u \in H_0^1(\Omega)$ 时满足, 即 $\int_\Omega \mathbf{grad}\, u \cdot \mathbf{grad}\, v dx = \int_\Omega f v dx$ 对所有 $v \in H_0^1(\Omega)$, 因此更不要说对所有 $v \in \mathcal{D}(\Omega)$ 了, 这证明了 $-\mathrm{div}\,\mathbf{grad}\, u = f \in L^2(\Omega)$. 所以 $\mathbf{grad}\, u \in \boldsymbol{H}(\mathrm{div};\Omega)$, 且

$$\int_\Omega (\mathrm{div}\,\mathbf{grad}\, u)\mu dx = -\int_\Omega f\mu dx \quad \text{对所有 } \mu \in L^2(\Omega).$$

因此, (b) 中变分问题的第二个方程当 $\boldsymbol{p} = \mathbf{grad}\, u$ 时满足. 这就证明了 (b).

因为集合 $\widetilde{\boldsymbol{U}}_f$ 可以等价地定义为

$$\widetilde{\boldsymbol{U}}_f = \{\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div}; \Omega); \widetilde{b}(\boldsymbol{q}, \mu) = \chi(\mu) \text{ 对所有 } \mu \in L^2(\Omega)\},$$

由于双线性形式 $a(\cdot, \cdot)$ 是对称的, 定理 6.12-2 适用. 这就证明了 (c). □

注 虽然出现在定理 6.13-2(c) 中的集合 $\widetilde{\boldsymbol{U}}_f$ 由所有满足

$$\mathrm{div}\, \boldsymbol{q} + f = 0 \quad \text{在 } L^2(\Omega) \text{ 中}$$

的向量场 $\boldsymbol{q} \in \boldsymbol{H}(\mathrm{div}; \Omega)$ 组成. 但前面已经说过, 出现在定理 6.13-1(c) 里集合 $\boldsymbol{U}_f$ 中的向量场 $\boldsymbol{q} \in \boldsymbol{L}^2(\Omega)$ 在 $H^{-1}(\Omega)$ 中满足同一个偏微分方程, 所以只在分布意义下满足.

本节对一个简单的模型问题所进行的分析讨论显然可以推广到定理 6.7-6 中所考虑的更一般的二阶椭圆边值问题, 即

$$-\sum_{i,j=1}^{N} \partial_i(a_{ij}\partial_j u) = f \text{ 在 } \Omega \text{ 中}, \quad u = 0 \text{ 在 } \partial\Omega \text{ 上}$$

(在这种情况下, 向量场 $(\sum_{j=1}^{N} a_{ij}\partial_j u)_{i=1}^{N} \in \boldsymbol{L}^2(\Omega)$ 起着模型问题中 $\mathbf{grad}\, u \in \boldsymbol{L}^2(\Omega)$ 的作用). 这种分析也可以推广到线性偏微分方程组的情形, 如 Stokes 方程 (6.14 节) 或线性化弹性方程组 (习题 6.16-3).

这类问题的混合及对偶形式具有重要意义, 因为它们是高效混合有限元法的基础[55].

6.14 Babuška-Brezzi 上下确界定理及 J. L. Lions 引理的应用: Stokes 方程组

本节的目标是确立 Stokes 方程组的存在性定理, 这个方程组是不可压缩黏性流体最常见的线性模型. 为此目的, 我们将验证 (定理 6.14-3), 这个方程组的弱形式构成曾在 6.12 节中予以讨论分析的抽象变分问题的另一个实例. 可以想见, 在存在性证明中, 关键的一步是验证 Babuška-Brezzi 上下确界条件 (定理 6.12-1) 成立. 这个验证依赖就其自身来说也是很重要的一个预备性结果, 此结果首先在定理 6.14-1 中给出.

值域在 $\mathbb{R}^N$ 中的向量场空间仍由粗体字母表示. 例如, $\boldsymbol{H}_0^1(\Omega)$ 表示所有向量场 $\boldsymbol{v} = (v_i)_{i=1}^{N}$ 的空间, 其分量 v_i 在空间 $H_0^1(\Omega)$ 中.

---

[55] 混合有限元法的详尽分析读者可以在下述著述中找到: Girault & Raviart [1986], Brezzi & Fortin [1991] 和 Roberts & Thomas [1991].

在整个这一节里, Ω 表示 $\mathbb{R}^N$ 中的区域, $\Gamma := \partial\Omega$. 作为开始, 我们引入一些函数空间及算子. Hilbert 空间 $\boldsymbol{H}_0^1(\Omega)$ 装备以如下定义的内积及范数:

$$(\boldsymbol{u},\boldsymbol{v})_{1,\Omega} := \int_\Omega \boldsymbol{\nabla u}\cdot\boldsymbol{\nabla u}dx \quad \text{对每组 } \boldsymbol{u},\boldsymbol{v}\in \boldsymbol{H}_0^1(\Omega),$$

其中

$$\boldsymbol{\nabla u}\cdot\boldsymbol{\nabla v} := \sum_{i,j=1}^N \partial_j u_i \partial_j v_i,$$

$$|\boldsymbol{v}|_{1,\Omega} := \sqrt{(\boldsymbol{v},\boldsymbol{v})_{1,\Omega}} \quad \text{对每个 } \boldsymbol{v}\in \boldsymbol{H}_0^1(\Omega),$$

而 Hilbert 空间

$$L_0^2(\Omega) := \left\{\mu\in L^2(\Omega); \int_\Omega \mu dx = 0\right\}$$

则装备以通常的空间 $L^2(\Omega)$ 的内积和范数, 它们分别以 $(\cdot,\cdot)_{0,\Omega}$ 和 $\|\cdot\|_{0,\Omega}$ 表示.

首先注意, $\int_\Omega \operatorname{div}\boldsymbol{v}dx = 0$ 对所有 $\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)$ 成立, 这是因为 $\operatorname{div}:\boldsymbol{H}_0^1(\Omega)\to \boldsymbol{L}^2(\Omega)$ 是一个连续算子 (这个关系式显然对分量 v_i 属于 $\mathcal{D}(\Omega)$ 的所有向量场 $\boldsymbol{v} = (v_i)_{i=1}^N$ 成立), 而 $\mathcal{D}(\Omega)$ 又是 $\boldsymbol{H}_0^1(\Omega)$ 的稠密子空间. 所以, 用这种方式定义的映射

$$\operatorname{div}:\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)\to \operatorname{div}\boldsymbol{v} := \sum_{i=1}^N \partial_i v_i \in L_0^2(\Omega)$$

是连续线性算子.

可以看出, 要证明出现在 Stokes 方程变分形式 (定理 6.14-3) 中的双线性形式 $b(\cdot,\cdot)$ 满足 Babuška-Brezzi 上下确界条件, 关键在于连续线性算子

$$\operatorname{div}:\boldsymbol{H}_0^1(\Omega)\to L_0^2(\Omega)$$

是满射, 而这正是下面定理 6.14-1 所要确立的一条性质.

这条看似无关痛痒的性质之证明却绝非平凡. 下面给出的证明特别依赖 J. L. Lions 引理 (定理 6.11-4) 和 Banach 闭值域定理 (定理 5.11-5)[56].

空间

$$\mathbf{Ker}\operatorname{div} := \{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega); \operatorname{div}v = 0 \text{ 在 } L^2(\Omega) \text{ 中}\}$$

是 $\boldsymbol{H}_0^1(\Omega)$ 的闭子空间, 直接和定理 (定理 4.5-2) 说明, 空间 $\boldsymbol{H}_0^1(\Omega)$ 可表示为

$$\boldsymbol{H}_0^1(\Omega) = \mathbf{Ker}\operatorname{div}\oplus(\mathbf{Ker}\operatorname{div})^\perp,$$

这里的正交性可理解为关于内积 $(\cdot,\cdot)_{1,\Omega}$. 下面将得 (显然是单射的) 算子 $\operatorname{div}: (\mathbf{Ker}\operatorname{div})^\perp\to L_0^2(\Omega)$ 有连续的逆, 这是因为它是满射 (见以下定理的 (c) 部分).

[56] $\operatorname{div}:\boldsymbol{H}_0^1(\Omega)\to L_0^2(\Omega)$ 是满射的第一个证明归于 LADYZHESKAYA [1969]; 也见于 TBMAM [1977, 第 1 章] 及 GIRAULT & RAVIART [1979, 3.3 节, 引理 3.2].

这个证明归之于引入映射

$$\mathbf{grad}: \mu \in L^2(\Omega) \to \mathbf{grad}\, \mu \in \boldsymbol{H}^{-1}(\Omega) := H^{-1}(\Omega; \mathbb{R}^N),$$

其定义由下式给出:

$$_{\boldsymbol{H}^{-1}(\Omega)}\langle \mathbf{grad}\, \mu, \boldsymbol{v} \rangle_{\boldsymbol{H}_0^1(\Omega)} := -\int_\Omega \mu \mathrm{div}\, \boldsymbol{v} dx \quad 对所有\ \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega).$$

如同每个映射 $\partial_i : L^2(\Omega) \to H^{-1}(\Omega), 1 \leqslant i \leqslant N$, 的定义, $\mathbf{grad} : L^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega)$ 在此也被定义在分布意义下. 像空间 $H^{-1}(\Omega)$ 那样, 空间 $\boldsymbol{H}^{-1}(\Omega)$ 作为空间 $\boldsymbol{H}_0^1(\Omega)$ 的对偶空间, 其范数用 $\|\cdot\|_{-1,\Omega}$ 表示.

所以这样定义的映射 $\mathbf{grad} : L^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega)$ 显然是线性而且连续的 (每一个映射 $\mu \in L^2(\Omega) \to \partial_i \mu \in H^{-1}(\Omega), 1 \leqslant i \leqslant N$, 都是连续的). 如下面证明中所显示, 当算子 $\mathbf{grad}$ 被限制于 $L^2(\Omega)$ 的子空间 $L_0^2(\Omega)$ 上时, 它是单射. 更为重要的是, 它其实是前面引入的算子 div 的对偶算子 (5.11 节).

要注意, 在定理 6.14-1 和定理 6.14-2 及其证明中, 我们默认空间 $L_0^2(\Omega)$ 与其对偶等同 (由于 $L_0^2(\Omega)$ 是 Hilbert 空间, 根据从 $(L_0^2(\Omega))'$ 到 $L_0^2(\Omega)$ 上的 F. Riesz 等距, 这是合理的), 也默认空间 $\boldsymbol{H}_0^1(\Omega)$ 与其双对偶空间等同 (由于 $\boldsymbol{H}_0^1(\Omega)$ 是自反的, 根据从 $\boldsymbol{H}_0^1(\Omega)$ 到 $(\boldsymbol{H}_0^1(\Omega))''$ 上的标准等距, 这也是合理的; 见 5.14 节).

定理 6.14-1　设 Ω 为 $\mathbb{R}^N$ 中的区域. 则

(a) 连续线性算子

$$\mathbf{grad} : L_0^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega)$$

对每个 $\mu \in L_0^2(\Omega)$ 由下式定义:

$$_{\boldsymbol{H}^{-1}(\Omega)}\langle \mathbf{grad} \mu, \boldsymbol{v} \rangle_{\boldsymbol{H}_0^1(\Omega)} := -\int_\Omega \mu \mathrm{div}\, \boldsymbol{v} dx \quad 对所有\ \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega).$$

该算子是单射, 且其对偶是连续线性算子

$$-\mathrm{div} : \boldsymbol{H}_0^1(\Omega) \to L_0^2(\Omega).$$

(b) 空间 $L_0^2(\Omega)$ 在算子 $\mathbf{grad}$ 作用下的像在 $\boldsymbol{H}^{-1}(\Omega)$ 中是闭的.

(c) 单射连续线性算子

$$\mathrm{div} : (\mathbf{Ker}\, \mathrm{div})^\perp \to L_0^2(\Omega)$$

是满射并有连续的逆算子.

证明　在下面的证明中, 同一个字母 C 表示各种常数, 这些常数在各种不同的场合可以不相等.

(i) 算子 $\mathbf{grad} : L_0^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega)$ 是单射.

设 $\mu \in L^2(\Omega)$ 使得 $\mathbf{grad}\,\mu = 0$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中. 由算子 $\mathbf{grad}$ 的定义, 这意味着

$$\sum_{i=1}^{N} \int_{\Omega} \mu \partial_i \varphi_i dx = 0 \quad \text{对所有 } \varphi_i \in \mathcal{D}(\Omega), \quad 1 \leqslant i \leqslant N.$$

故由定理 6.3-4 得 μ 为常数, 因此 $\mu = 0$ 若 $\mu \in L_0^2(\Omega)$.

(ii) 算子 $-\mathrm{div} : \boldsymbol{H}_0^1(\Omega) \to L_0^2(\Omega)$ 是算子 $\mathbf{grad}$: $L_0^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega)$ 的对偶 (即在赋范向量空间的意义下, 见 5.11 节). 令 $\sigma : \boldsymbol{H}^{-1}(\Omega) \to \boldsymbol{H}_0^1(\Omega)$ 表示 Hilbert 空间 $\boldsymbol{H}_0^1(\Omega)$ 的 F. Riesz 等距, 则算子 $\sigma\mathbf{grad} : L_0^2(\Omega) \to \boldsymbol{H}_0^1(\Omega)$ 是 $-\mathrm{div} : \boldsymbol{H}_0^1(\Omega) \to L_0^2(\Omega)$ 的伴随算子 (即在 Hilbert 空间的意义下; 见 4.7 节).

关系式

$${}_{\boldsymbol{H}^{-1}(\Omega)}\langle \mathbf{grad}\,\mu, \boldsymbol{v} \rangle_{\boldsymbol{H}_0^1(\Omega)} = -\int_{\Omega} \mu \mathrm{div}\, \boldsymbol{v} dx \quad \text{对所有 } \mu \in L_0^2(\Omega) \text{ 和所有 } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$$

说明 $-\mathrm{div} : \boldsymbol{H}_0^1(\Omega) \to L_0^2(\Omega)$ 是 $\mathbf{grad} : L_0^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega)$ 的对偶 (在赋范向量空间的意义下). 用 F. Riesz 等距 $\sigma : \boldsymbol{H}^{-1}(\Omega) \to \boldsymbol{H}_0^1(\Omega)$ 来表示, 同一关系式变为

$$(\sigma\mathbf{grad}\,\mu, \boldsymbol{v})_{1,\Omega} = -(\mu, \mathrm{div}\, \boldsymbol{v})_{0,\Omega} \quad \text{对所有 } \mu \in L_0^2(\Omega) \text{ 和所有 } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega).$$

所以这证明了 $-\mathrm{div} : \boldsymbol{H}_0^1(\Omega) \to L_0^2(\Omega)$ 是 $\sigma\mathbf{grad} : L_0^2(\Omega) \to \boldsymbol{H}_0^1(\Omega)$ 的伴随算子 (在 Hilbert 空间的意义下).

(iii) *存在常数 C 使得*[57)]

$$\|\mu\|_{0,\Omega} \leqslant C(\|\mu\|_{-1,\Omega}^2 + \|\mathbf{grad}\,\mu\|_{-1,\Omega}^2)^{1/2} \quad \text{对所有 } \mu \in L^2(\Omega).$$

首先我们断言, 空间

$$K(\Omega) := \{\mu \in H^{-1}(\Omega); \mathbf{grad}\,\mu \in \boldsymbol{H}^{-1}(\Omega)\}$$

(在这个定义中, $\mathbf{grad}$ 仍理解为在分布意义下), 装备以范数

$$\mu \in K(\Omega) \to \|\mu\|_{K(\Omega)} := (\|\mu\|_{-1,\Omega}^2 + \|\mathbf{grad}\,\mu\|_{-1,\Omega}^2)^{1/2}$$

是完备的. 为证明这一断言, 设 $(\mu_k)_{k=1}^{\infty}$ 是 $K(\Omega)$ 中的 Cauchy 序列. 则有 $\mu_k \xrightarrow[k\to\infty]{} \mu$ 在 $H^{-1}(\Omega)$ 中及 $\mathbf{grad}\,\mu_k \xrightarrow[k\to\infty]{} \boldsymbol{w}$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中 (作为对偶空间, $H^{-1}(\Omega)$ 及 $\boldsymbol{H}^{-1}(\Omega)$ 是完备的), 且

$${}_{\boldsymbol{H}^{-1}(\Omega)}\langle \mathbf{grad}\,\mu_k, \boldsymbol{\varphi} \rangle_{\boldsymbol{H}_0^1(\Omega)} = -\boldsymbol{H}^{-1}(\Omega)\langle \mu_k, \mathrm{div}\boldsymbol{\varphi} \rangle_{\boldsymbol{H}_0'(\Omega)} \quad \text{对所有 } \boldsymbol{\varphi} \in \mathcal{D}(\Omega; \mathbb{R}^N).$$

在以上关系式中取极限, 令 $k \to \infty$ 即得

$${}_{\boldsymbol{H}^{-1}(\Omega)}\langle \boldsymbol{w}, \varphi \rangle_{\boldsymbol{H}_0^1(\Omega)} = -\boldsymbol{H}^{-1}(\Omega)\langle \mu, \mathrm{div}\varphi \rangle_{\boldsymbol{H}_0'(\Omega)} \quad \text{对所有 } \boldsymbol{\varphi} \in \mathcal{D}(\Omega; \mathbb{R}^N).$$

---

57) 这个关键的不等式的直接尽管微妙的证明见 NEČAS [1965].

这就证明了 $\boldsymbol{w} = \mathbf{grad}\,\mu$. 所以 $(K(\Omega), \|\cdot\|_{K(\Omega)})$ 是完备的.

典范内射 $\iota : (L^2(\Omega), \|\cdot\|_{0,\Omega}) \to (K(\Omega), \|\cdot\|_{K(\Omega)})$ 是连续的 (显然存在常数 C 使得 $\|\mu\|_{K(\Omega)} \leqslant C\|\mu\|_{0,\Omega}$ 对所有 $\mu \in L^2(\Omega)$) 单射, 而根据 J. L. Lions 引理 (定理 6.11-4) $K(\Omega) = L^2(\Omega)$, 故又是满射.

所以, 由 Banach 开映射定理的推论 (定理 5.6-2) 知, 逆映射 ι^{-1} 也是连续的, 故存在常数 C 使得

$$\|\mu\|_{0,\Omega} \leqslant C(\|\mu\|^2_{-1,\Omega} + \|\mathbf{grad}\,\mu\|^2_{-1,\Omega})^{1/2} \quad 对所有\ \mu \in L^2(\Omega).$$

(iv) 存在常数 C 使得

$$\|\mu\|_{0,\Omega} \leqslant C\|\mathbf{grad}\,\mu\|_{-1,\Omega} \quad 对所有\ \mu \in L^2_0(\Omega).$$

我们用反证法. 如果上式不成立, 存在一个函数 $\mu_k \in L^2_0(\Omega)$ 的序列 $(\mu_k)_{k=1}^\infty$ 使得

$$\|\mu_k\|_{0,\Omega} = 1 \quad 对所有\ k \geqslant 1 \quad 且 \quad \|\mathbf{grad}\,\mu_k\|_{-1,\Omega} \to 0 \quad 当\ k \to \infty\ 时.$$

根据在空间 $L^2(\Omega)$ 中的 Rellich-Kondrachov 紧嵌入定理 (定理 6.11-3), 存在一个在 $H^{-1}(\Omega)$ 中收敛的子序列 $(\mu_{\sigma(k)})_{k=1}^\infty$. 由于子序列 $(\mathbf{grad}\,\mu_{\sigma(k)})_{k=1}^\infty$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中收敛 (于 0, 但现阶段并不需要用到这一事实), 因此子序列 $(\mu_{\sigma(k)})_{k=1}^\infty$ 是空间 $(K(\Omega), \|\cdot\|_{K(\Omega)})$ 中的 Cauchy 序列, 由 (iii) 知也是空间 $L^2(\Omega)$ 中的 Cauchy 序列.

设 $\mu \in L^2(\Omega)$ 使得

$$\mu_{\sigma(k)} \xrightarrow[k\to\infty]{} \mu \quad 在\ L^2(\Omega)\ 中.$$

则 $\mu \in L^2_0(\Omega)$, 此因 $\displaystyle\int_\Omega \mu dx = \lim_{k\to\infty} \int_\Omega \mu_{\sigma(k)} dx = 0$; 又有 $\mathbf{grad}\,\mu = 0$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中, 此因

$$\mathbf{grad}\,\mu_\sigma(k) \xrightarrow[k\to\infty]{} \mathbf{0} = \mathbf{grad}\,\mu \quad 在\ \boldsymbol{H}^{-1}(\Omega)\ 中,$$

故由定理 6.3-4 得 $\mu = 0$ ($\mathbf{grad}\,\mu = 0$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中意味着 $\displaystyle\int_\Omega \mu\partial_i\varphi dx = 0$ 对所有 $\varphi \in \mathcal{D}(\Omega), 1 \leqslant i \leqslant N$, 故 μ 为常数, 但 $\mu \in L^2_0(\Omega)$, 所以这个常数为零). 这与关系式 $\|\mu_{\sigma(k)}\|_{0,\Omega} = 1$ 对所有 $k \geqslant 1$ 矛盾.

(v) $L^2_0(\Omega)$ 在 $\mathbf{grad}$ 作用下的像在 $\boldsymbol{H}^{-1}(\Omega)$ 中是闭的, $\boldsymbol{H}^1_0(\Omega)$ 在 div 作用下的像在 $L^2_0(\Omega)$ 中是闭的.

在 (iv) 中已证明, 存在常数 C 使得

$$\|\mu\|_{0,\Omega} \leqslant C\|\mathbf{grad}\,\mu\|_{-1,\Omega} \quad 对所有\ \mu \in L^2_0(\Omega).$$

由于 $L^2_0(\Omega)$ 是完备的 (定理 3.1-4), 所以 $L^2_0(\Omega)$ 在 $\mathbf{grad}$ 作用下的像在 $\boldsymbol{H}^{-1}(\Omega)$ 中是完备的.

因为 $-\mathrm{div}:\boldsymbol{H}_0^1(\Omega)\to L_0^2(\Omega)$ 是 $\mathbf{grad}:L_0^2(\Omega)\to\boldsymbol{H}^{-1}(\Omega)$ 的对偶算子 (见 (ii)), 由 Banach 闭值域定理 (第一部分, 见定理 5.11-5) 知 $\boldsymbol{H}_0^1(\Omega)$ 在 div 作用下的像在 $L_0^2(\Omega)$ 中也是闭的.

(vi) 单射算子 $\mathrm{div}:(\mathbf{Ker}\ \mathrm{div})^{\perp}\to L_0^2(\Omega)$ 是满射且有连续的逆.

(v) 实际上也证明了 $L_0^2(\Omega)$ 在算子 $\sigma\mathbf{grad}:L_0^2(\Omega)\to\boldsymbol{H}_0^1(\Omega)$ 作用下的像在 $\boldsymbol{H}_0^1(\Omega)$ 中是闭的, 这是因为 F. Riesz 映射 $\sigma:\boldsymbol{H}^{-1}(\Omega)\to\boldsymbol{H}_0^1(\Omega)$ 是一个等距. 另一方面定理 4.7-2(b) 说明

$$L_0^2(\Omega)=\mathrm{Ker}(\sigma\mathbf{grad})\oplus\mathrm{Im}(-\mathrm{div}).$$

所以算子 $\mathrm{div}:\boldsymbol{H}_0^1(\Omega)\to L_0^2(\Omega)$ 是满射, 这是因为算子 $\sigma\mathbf{grad}:L_0^2(\Omega)\to\boldsymbol{H}_0^1(\Omega)$ 是单射 (如同算子 $\mathbf{grad}$, 见 (i)), 故 $\mathrm{Ker}(\sigma\mathbf{grad})=\{0\}$. 因此 $\mathrm{div}:(\mathbf{Ker}\ \mathrm{div})^{\perp}\to L_0^2(\Omega)$ 是双射, 根据 Banach 开映射定理的推论, 其逆也是连续的. 所以 (c) 得证. □

注 在习题 6.14-1 到 6.14-3 中, 给出了关于定理 6.14-1 有趣的补充材料.

作为定理 6.14-1 的一个推论, 我们下面确立 $\boldsymbol{H}^{-1}(\Omega)$ 中的一个向量场成为 $L^2(\Omega)$ 中数量函数梯度的第一个表征; 弱 Poincaré 引理 (稍后给出; 见定理 6.17-4) 则成为这种向量场 (在 Ω 是单连通这一附加假设下) 的第二个表征.

定理 6.14-2 设 Ω 是 $\mathbb{R}^N$ 中的区域. 给定一个向量场 $\boldsymbol{h}\in\boldsymbol{H}^{-1}(\Omega)$, 存在一个函数 p 使得

$$p\in L_0^2(\Omega)\quad 且\quad \mathbf{grad}\,p=\boldsymbol{h}\ 在\ \boldsymbol{H}^{-1}(\Omega)\ 中$$

的充分必要条件为

$${}_{\boldsymbol{H}^{-1}(\Omega)}\langle\boldsymbol{h},\boldsymbol{v}\rangle_{\boldsymbol{H}_0^1(\Omega)}=0\quad 对在\ L^2(\Omega)\ 中满足\ \mathrm{div}\,\boldsymbol{v}=0\ 的所有\ \boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega).$$

在 $\boldsymbol{H}^{-1}(\Omega)$ 中满足方程 $\mathbf{grad}\,\tilde{p}=\boldsymbol{h}$ 的所有其他的解 $\tilde{p}\in L^2(\Omega)$ 均形如 $\tilde{p}=p+C$, 其中 C 是常数.

证明 因为 $\mathbf{grad}:L_0^2(\Omega)\to\boldsymbol{H}^{-1}(\Omega)$ 的对偶是 $-\mathrm{div}:H_0^1(\Omega)\to L_0^2(\Omega)$, 而且 $L_0^2(\Omega)$ 在 $\mathbf{grad}$ 的作用下的像 $\mathbf{Im\ grad}$ 于 $\boldsymbol{H}^{-1}(\Omega)$ 中是闭的 (定理 6.14-1), 根据 Banach 闭值域定理 (见定理 5.11-5, 第一部分) 有

$$\mathbf{Im\ grad}=\{\boldsymbol{h}\in\boldsymbol{H}^{-1}(\Omega);\,{}_{\boldsymbol{H}^{-1}(\Omega)}\langle\boldsymbol{h},\boldsymbol{v}\rangle_{\boldsymbol{H}_0^1(\Omega)}=0\ 对所有\ \boldsymbol{v}\in\mathbf{Ker}(-\mathrm{div})\}.$$

换言之, 正如定理中所称, 给定 $\boldsymbol{h}\in\boldsymbol{H}^{-1}(\Omega)$, 存在方程 $\mathbf{grad}\,p=\boldsymbol{h}$ 的一个解 $p\in L_0^2(\Omega)\subset L^2(\Omega)$ 的充分必要条件是 ${}_{\boldsymbol{H}^{-1}(\Omega)}\langle\boldsymbol{h},\boldsymbol{v}\rangle_{\boldsymbol{H}_0^1(\Omega)}=0$ 对所有在 $L^2(\Omega)$ 中满足 $\mathrm{div}\,\boldsymbol{v}=0$ 的 $\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)$ 成立.

设 $\pi\in L^2(\Omega)$ 满足 $\mathbf{grad}\,\pi=\mathbf{0}$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中; 特别地有

$${}_{\boldsymbol{H}^{-1}(\Omega)}\langle\partial_i\pi,\varphi\rangle_{H_0^1(\Omega)}:=-\int_\Omega\pi\partial_i\varphi dx=0\quad 对所有\ \varphi\in\mathcal{D}(\Omega),\quad 1\leqslant i\leqslant N.$$

故由定理 6.3-4 (按假设, 区域是连通的), 函数 π 是常数. 这证明了, $\mathbf{grad}\,\tilde{p} = \boldsymbol{h}$ 的所有其他解形如 $\tilde{p} = p + C$, 其中 C 为某常数. □

我们现在讨论 Stokes 方程组. 如同处理二阶线性椭圆边值问题那样 (6.7 节), 首先给出一组特定的变分方程, 然后证明这组方程有唯一解, 最后验明其解相应的边值问题 (这里, 照例假定变分方程的解足够光滑). 由于其中一个未知量是向量场, 即 $\boldsymbol{u} = (u_i) \in \boldsymbol{H}_0^1(\Omega)$, 这个问题形成一个偏微分方程组 (而不是 6.7 节中那样的单个方程).

定理 6.14-3 (Stokes 方程组解的存在性) 设 Ω 是 $\mathbb{R}^N$ 中的区域, 而常数 $\nu > 0$ 及向量场 $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ 是给定的. 则

(a) 存在唯一的偶对 $(\boldsymbol{u}, \lambda) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ 满足变分问题

$$\nu \int_\Omega \boldsymbol{\nabla u} : \boldsymbol{\nabla v} dx - \int_\Omega (\operatorname{div} \boldsymbol{v}) \lambda dx = \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} dx \quad \text{对所有 } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega),$$
$$\int_\Omega (\operatorname{div} \boldsymbol{u}) \mu dx = 0 \quad \text{对所有 } \mu \in L_0^2(\Omega),$$

后一个方程等价于

$$\operatorname{div} \boldsymbol{u} = 0 \quad \text{在 } L_0^2(\Omega) \text{ 中}.$$

(b) 向量场 $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ 是以下约束二次极小化问题的唯一解:

$$\boldsymbol{u} \in \boldsymbol{U}_0 := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega); \operatorname{div} \boldsymbol{v} = 0 \text{ 在 } \Omega \text{ 中}\},$$
$$I(\boldsymbol{u}) = \inf_{\boldsymbol{v} \in \boldsymbol{U}_0} I(\boldsymbol{v}),$$

其中

$$I(\boldsymbol{v}) := \frac{\nu}{2} \int_\Omega \boldsymbol{\nabla u} : \boldsymbol{\nabla v} dx - \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} dx \quad \text{对每个 } \boldsymbol{v} \in \boldsymbol{U}_0.$$

(c) 偶对 $(\boldsymbol{u}, \lambda) \in \boldsymbol{H}_0^1(\Omega) \times L_0^2(\Omega)$ 满足

$$-\nu \Delta \boldsymbol{u} + \mathbf{grad}\, \lambda = \boldsymbol{f} \quad \text{在 } \boldsymbol{H}^{-1}(\Omega) \text{ 中},$$
$$\operatorname{div} \boldsymbol{u} = 0 \quad \text{在 } \Omega \text{ 中},$$
$$\boldsymbol{u} = \boldsymbol{0} \quad \text{在 } \Gamma \text{ 上},$$

其中 $\boldsymbol{\Delta u} := (\Delta u_i)_{i=1}^N$.

证明 一个向量场 $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ 满足出现在 (a) 中的关系式 $\operatorname{div} \boldsymbol{u} = 0$ 在 $L_0^2(\Omega)$ 中, 当 (且显然仅当)

$$\int_\Omega (\operatorname{div} \boldsymbol{u}) \mu dx = 0 \quad \text{对所有 } \mu \in L_0^2(\Omega)$$

成立 (为说明这一点, 可在上式中令 $\mu = \operatorname{div} \boldsymbol{u}$).

设双线性形式 $a(\cdot,\cdot):\boldsymbol{H}_0^1(\Omega)\times\boldsymbol{H}_0^1(\Omega)\to\mathbb{R}, b(\cdot,\cdot):\boldsymbol{H}_0^1(\Omega)\times L_0^2(\Omega)\to\mathbb{R}$ 及线性形式 $\ell:\boldsymbol{H}_0^1(\Omega)\to\mathbb{R},\chi:L_0^2(\Omega)\to\mathbb{R}$ 分别由以下诸式定义:

$$a(\boldsymbol{u},\boldsymbol{v}):=\nu\int_\Omega\boldsymbol{\nabla u}\cdot\boldsymbol{\nabla v}dx\quad 对每个\ \boldsymbol{u},\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega),$$
$$b(\boldsymbol{v},\mu):=\int_\Omega(\operatorname{div}\boldsymbol{v})\mu dx\quad 对每对\ (\boldsymbol{v},\mu)\in\boldsymbol{H}_0^1(\Omega)\times L_0^2(\Omega),$$
$$\ell(\boldsymbol{v}):=\int_\Omega\boldsymbol{f}\cdot\boldsymbol{v}dx\quad 对每个\ \boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)\ 及\ \chi:=0.$$

对称双线性形式 $a(\cdot,\cdot)$ 显然是连续的, 而且是 $\boldsymbol{H}_0^1(\Omega)$ 强制的, 此因

$$a(\boldsymbol{v},\boldsymbol{v})=\nu|\boldsymbol{v}|_{1,\Omega}^2\quad 对所有\ \boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega).$$

所以 $a(\cdot,\cdot)$ 更应在

$$\boldsymbol{U}_0=\{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega);b(\boldsymbol{v},\mu)=0\ 对所有\ \mu\in L_0^2(\Omega)\}$$

上强制.

根据定理 6.14-1(c), 给定任意函数 $\mu\in L_0^2(\Omega)$, 存在唯一的向量场 $w_\mu\in(\mathbf{Ker}\operatorname{div})^\perp\subset\boldsymbol{H}_0^1(\Omega)$ 满足

$$\operatorname{div}\boldsymbol{w}_\mu=\mu\quad 在\ L_0^2(\Omega)\ 中,$$

此外还存在常数 C 使得

$$|\boldsymbol{w}_\mu|_{1,\Omega}\leqslant C\|\mu\|_{0,\Omega}\quad 对所有\ \mu\in L_0^2(\Omega).$$

这样, 对每个非零的 $\mu\in L_0^2(\Omega)$, 有

$$\sup_{\left\{\substack{\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)\\ \boldsymbol{v}\neq\boldsymbol{0}}\right.}\frac{\left|\int_\Omega(\operatorname{div}\boldsymbol{v})\mu dx\right|}{|\boldsymbol{v}|_{1,\Omega}}\geqslant\frac{\int_\Omega(\operatorname{div}\boldsymbol{w}_\mu)\mu dx}{|\boldsymbol{w}_\mu|_{1,\Omega}}=\frac{\|\mu\|_{0,\Omega}^2}{|\boldsymbol{w}_\mu|_{1,\Omega}}\geqslant C^{-1}\|\mu\|_{0,\Omega},$$

这说明定理 6.12-1 中的 Babuška-Brezzi 上下确界条件成立, 其中

$$V:=\boldsymbol{H}_0^1(\Omega),\quad M:=L_0^2(\Omega).$$

线性形式 ℓ 在空间 $\boldsymbol{H}_0^1(\Omega)$ 上显然是连续的. 所以 (a) 与 (b) 可分别由定理 6.12-1 及定理 6.12-2 得到.

最后, 关系式

$$\nu\int_\Omega\sum_{i,j=1}^N\partial_ju_i\partial_jv_idx-\int_\Omega\lambda\sum_{i=1}^N\partial_iv_idx$$

$$= \sum_{i=1}^{N} {}_{H^{-1}(\Omega)}\langle -\nu\Delta u_i + \partial_i\lambda, v_i\rangle_{H_0^1(\Omega)}$$

$$= \int_\Omega \sum_{i=1}^{N} f_i v_i dx \quad \text{对所有 } (v_i)_{i=1}^N \in \boldsymbol{H}_0^1(\Omega)$$

证明了方程 $-\nu\Delta u_i + \partial_i\lambda = f_i$ 在 $H^{-1}(\Omega)$ 中, $1 \leqslant i \leqslant N$, 成立. □

以下诸式

$$-\nu\Delta\boldsymbol{u} + \mathbf{grad}\,\lambda = \boldsymbol{f} \quad \text{在 } \Omega \text{ 中},$$
$$\operatorname{div}\boldsymbol{u} = 0 \quad \text{在 } \Omega \text{ 中},$$
$$\boldsymbol{u} = \mathbf{0} \quad \text{在 } \Gamma \text{ 上},$$

构成 $\mathbb{R}^N$ 中的 Stokes *方程组*[58]. 当 $N = 3$ 时, 这组方程成为非线性 Navier-Stokes 方程组 (9.11 节) 的线性化, 该方程组是*黏性不可压缩流体*的定常 (即与时间无关) 流的模型, 流体充满区域 $\Omega \subset \mathbb{R}^3$, 具有*动力黏性* $\nu > 0$, 且单位质量流体受到密度为 $\boldsymbol{f}$ 的外力作用. 未知量 $\boldsymbol{u}$ 是流体的*速度*, 它在 Ω 中服从*不可压缩条件* $\operatorname{div}\boldsymbol{u} = 0$ 而在 Γ 上服从边界条件 $\boldsymbol{u} = \mathbf{0}$, 这意味着流体的速度在区域的边界上为零. 未知量 λ 表示流体内部的*压力*.

值得注意的是, 在作为约束二次极小化问题的 Stokes 问题的形式 (定理 6.14-3(b)) 中, 未知量 λ 并不出现, 而出现在其混合形式 (定理 6.14-3(a)) 中. 稍后 (7.16 节) 我们将看到, 未知量 λ 也出现在其作为*鞍点问题*以及作为与*约束* $\operatorname{div}\boldsymbol{v} = 0$ (在 Ω 内) *相应的* Lagrange *乘子*形式中.

要指出的是, 如果未知量 λ 不是被看作在空间 $L_0^2(\Omega)$ 内, 而是视为在商空间 $L^2(\Omega)/P_0(\Omega)$ 中, 其中 $P_0(\Omega)$ 表示 Ω 上常函数组成的空间 (容易看出, 在 $L_0^2(\Omega)$ 与 $L^2(\Omega)/P_0(\Omega)$ 之间存在线性等距), 对于 Stokes 方程组的数学分析可以同样地进行. 特别地, 这种考虑反映出, *未知压力 λ 被唯一确定只相差一个附加常数* (如果我们的出发点是上面的边值问题, 这一性质是显然的). 如果选取的空间是 $L_0^2(\Omega)$, 未知量 λ 必须满足条件 $\int_\Omega \lambda dx = 0$, 当然就不会存在这种不确定性.

在整个的边界 Γ 上给出的未知速度 $\boldsymbol{u}$ 满足边界条件 $\boldsymbol{u} = \mathbf{0}$, 尽管可在整个边界 Γ 上给出更一般的边界条件 $\boldsymbol{u} = \boldsymbol{u}_0$, 但大家普遍认为这是不符合实际的. 然而, 考虑物理上更合理的边界条件 (例如, 自由表面边界条件) 在数学上有相当的难度. 这就是更符合实际的边界条件很少予以讨论的原因[59].

---

[58] 冠名源自 Sir George Gabriel Stokes (1819—1903).

[59] 一个值得关注的例外, 可见于下文中:

V. A. SOLONNIKOV [1982]: On the Stokes equations in domains with non-smooth boundaries and on viscous incompressible flow with a free surface. Nonlinear Partial Differential Equations and Their Applications (H. BREZIS & J. L. LIONS, editors), pp. 340–423, Pitman, Boston.

习题

6.14-1 [60] 给定 $\mathbb{R}^N$ 中的区域 Ω, 定义空间

$$\boldsymbol{V}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega); \operatorname{div} \boldsymbol{v} = 0 \text{ 在 } \Omega \text{ 内}\},$$

$$\mathcal{V}(\Omega) := \{\boldsymbol{\varphi} \in \mathcal{D}(\Omega; \mathbb{R}^N); \operatorname{div} \boldsymbol{\varphi} = 0 \text{ 在 } \Omega \text{ 内}\}.$$

(1) 设向量场 $\boldsymbol{h} \in \boldsymbol{H}^{-1}(\Omega)$ 使得 ${}_{\boldsymbol{H}^{-1}(\Omega)}\langle \boldsymbol{h}, \boldsymbol{v}\rangle_{\boldsymbol{H}_0^1(\Omega)} = 0$ 对所有 $\boldsymbol{\varphi} \in \mathcal{V}(\Omega)$. 证明存在一个函数 $p \in L^2(\Omega)$ 使得 $\boldsymbol{h} = \mathbf{grad}\, p$.

提示: 因为 Ω 是区域, 应是连通的. 故存在 $\mathbb{R}^N$ 中连通区域序列 $(\Omega_k)_{k=1}^{\infty}$, 具有以下性质:

$$\Omega_k \subset \Omega_{k+1}, \overline{\Omega}_k \subset \Omega \quad \text{对所有 } k \geqslant 1, \text{ 且 } \Omega = \bigcup_{k=1}^{\infty} \Omega_k.$$

给定任意的 $\boldsymbol{v} = (v_i)_{i=1}^N \in \boldsymbol{H}_0^1(\Omega)$ 及任意的 $\varepsilon > 0$, 设 $\boldsymbol{v}_\varepsilon := (v_{i,\varepsilon})_{i=1}^N$, 其中 $(v_{i,\varepsilon})_{\varepsilon>0}$ 是函数 $v_i \in H_0^1(\Omega), 1 \leqslant i \leqslant N$, 的正则化族 (2.6 节). 证明, 对于每个 $k \geqslant 1$, 存在 $\varepsilon(k) > 0$ 使得, 给定任意在 $\Omega - \Omega_k$ 上满足 $\boldsymbol{v} = \mathbf{0}$ 的 $\boldsymbol{v} \in \boldsymbol{V}(\Omega)$, 则 $\boldsymbol{v}_\varepsilon \in \boldsymbol{\nu}(\Omega)$ 对所有 $0 < \varepsilon \leqslant \varepsilon(k)$ 且 $|\boldsymbol{v}_\varepsilon - \boldsymbol{v}|_{1,\Omega} \to 0$ 当 $\varepsilon \to 0$ 时. 由这一结果推得 ${}_{\boldsymbol{H}^{-1}(\Omega)}\langle \boldsymbol{h}, \boldsymbol{v}\rangle_{\boldsymbol{H}_0^1(\Omega)} = 0$; 然后用定理 6.14-2.

注 (1) 中证明的结果是 de Rham 定理[61] 的一个特殊情况, 其更一般的深刻结果断言, 给定 $\mathbb{R}^N$ 中任意开子集 Ω, Ω 上任一向量值分布如果在空间 $\mathcal{V}(\Omega)$ 上为零一定是 Ω 上一分布的梯度.

(2) 利用 (1) 和定理 4.3-2 证明, $\boldsymbol{V}(\Omega)$ 的子空间 $\mathcal{V}(\Omega)$ 在空间 $(\boldsymbol{V}(\Omega), |\cdot|_{1,\Omega})$ 中稠密.

6.14-2 设 Ω 是 $\mathbb{R}^N$ 中的区域. 将算子 $\mathbf{grad}$ 视为从 $H_0^1(\Omega)$ 作用到 $\boldsymbol{L}^2(\Omega)$ (而不是如在定理 6.14-1 中那样从 $L_0^2(\Omega)$ 到 $\boldsymbol{H}^{-1}(\Omega)$), 本题列出该算子的两条性质.

(1) 证明 $-\operatorname{div} : \boldsymbol{L}^2(\Omega) \to H^{-1}(\Omega)$ 是算子 $\mathbf{grad} : H_0^1(\Omega) \to \boldsymbol{L}^2(\Omega)$ 的对偶算子.

(2) 证明空间 $H_0^1(\Omega)$ 在算子 $\mathbf{grad}$ 作用下的像在 $\boldsymbol{L}^2(\Omega)$ 中是闭的 (其证明较之定理 6.14-1 的 (b) 部分要简单得多, 因为在此不再依赖 J. L. Lions 引理).

6.14-3 给定任意的 $\boldsymbol{\ell} \in \boldsymbol{H}^{-1}(\Omega)$, 设 $\boldsymbol{u} := \boldsymbol{\mathcal{A}}(\boldsymbol{\ell}) \in \boldsymbol{H}_0^1(\Omega)$ 表示变分方程 $(\boldsymbol{u}, \boldsymbol{v})_{1,\Omega} = \boldsymbol{\ell}(\boldsymbol{v})$ 对所有 $\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega)$ 的唯一解. 证明

$$(\mathbf{Ker}\ \operatorname{div})^{\perp} = \{\mathcal{A}(\mathbf{grad}\, \mu) \in \boldsymbol{H}_0^1(\Omega); \mu \in L^2(\Omega)\},$$

其中映射 div 从 $\boldsymbol{H}_0^1(\Omega)$ 作用到 $L^2(\Omega)$ 中.

提示: 利用习题 6.14-1 中的问题 (1).

6.14-4 设 Ω 为 $\mathbb{R}^N$ 中的区域. 证明, 空间 $\boldsymbol{V}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega); \operatorname{div} \boldsymbol{v} = 0 \text{ 在 } \Omega \text{ 中}\}$ 关于范数 $\|\cdot\|_{0,\Omega}$ 的闭包是空间 $\{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega); \operatorname{div} \boldsymbol{v} = 0 \text{ 在 } H^{-1}(\Omega) \text{ 中}\}$ 的严格子空间 (当然, 与空间 $\boldsymbol{V}(\Omega)$ 一样, 同样的性质对于出现在习题 6.14-1 中的空间 $\mathcal{V}(\Omega)$ 也成立).

---

[60] 这个习题的问题 (1) 及 (2) 分别构成 GIRAULT & RAVIART [1986, 第 1 章] 的定理 2.3 及推论 2.5.

[61] G. de RHAM [1955]: Variétés Différentiables. Herwann, Paris.

6.15 J. L. Lions 引理的第二个应用: Korn 不等式

J. L. Lions 引理的第二个应用是证明一个基本不等式, 这一不等式在线性化弹性理论中起着关键的作用.

Korn 不等式[62] 断言, 给定 $\mathbb{R}^N$ 中一区域 Ω, 存在一个只依赖 Ω 的常数 C 使得

$$\left(\sum_{i=1}^N \|v_i\|_{0,\Omega}^2 + \sum_{i,j=1}^N \|\partial_j v_i\|_{0,\Omega}^2\right)^{1/2} \leqslant C \left(\sum_{i=1}^N \|v_i\|_{0,\Omega}^2 + \sum_{i,j=1}^N \|e_{ij}(v)\|_{0,\Omega}^2\right)^{1/2}$$

对所有向量场 $\boldsymbol{v} = (v_i)_{i=1}^N \in H^1(\Omega;\mathbb{R}^N)$, 其中

$$e_{ij}(\boldsymbol{v}) := \frac{1}{2}(\partial_j v_i + \partial_i v_j) \in L^2(\Omega), \quad 1 \leqslant i,j \leqslant N.$$

在下一节 (定理 6.16-1) 中我们将看到, 这个不等式 $N=3$ 的特殊情况, 在确立三维线性化弹性边值问题弱形式解的存在唯一性 (关键是证明相应的双线性形式的强制性) 中起着决定性的作用.

Korn 不等式为一个向量场 $\boldsymbol{v} = (v_i) \in \boldsymbol{H}^1(\Omega;\mathbb{R}^N)$ 所有 N^2 个偏导数 $\partial_j v_i$ 的 $L^2(\Omega)$ 范数提供了一个上界, 这个上界是只用了这些偏导数的 $\dfrac{N(N+1)}{2}$ 个特殊的线性组合, 即函数 $e_{ij}(\boldsymbol{v}) = e_{ji}(\boldsymbol{v})$ 的 $L^2(\Omega)$ 范数给出的. 这一异乎寻常的特性揭示出, 在该不等式的各种有效证明中[63], 没有哪一个会是简单的. 例如, 下面给出的

---

[62] 这个不等式起初是在向量场 $\boldsymbol{v}$ 在 Ω 的边界上为零的假设下予以证明的, 见下述诸文:

A. KORN [1906]: Die Eigenschwingungen eines elastischen Körpers mit ruhender Oberfläche. Sitzungsberichte der Mathematisch-physikalischen Klass der Königlich bayerischen Akademie der Wissenschaften zu München **36**, 351–402.

A. KORN [1908]: Solution générale du problème d'équilibre dans la théorie de l'élasticité. le cas où les efforts sont donnés à la surface, Annales de la Faculté des Sciences de Toulouse **10**, 165–269.

A. KORN [1909]: Über einige Ungleichungen, welche in der Theorie der elastischen und elektrischen Schwingungen eine Rolle spielen. Bulletin International de l'Académie des Sciences de Cracovie **9**, 705–724.

第二个证明是在向量场 $\boldsymbol{v}$ 满足 $\int_\Omega \mathbf{curl}\, \boldsymbol{v} dx = 0$ 的假设下给出的, 发表于:

K. O. FRIEDRICHS [1947]: On the boundary-value problems of the theory of elasticity and Korn's inequality. Annals of Mathematics **48**, 441–471.

完全一般情况下的第一个证明 (基于 Calderón-Zygmund 奇异积分理论) 归之于:

J. GOBERT [1962]: Une inégalité fondamentale de la théorie de l'élasticité. Bulletin de la Société Royale des Sciences de Liége **31**, 182–191.

[63] 例如, 可见下文所列的参考文献:

C. O. HORGAN [1995]: Kornś inequalities and their applications in continuum mechanics, SIAM Review **37**, 491–511.

证明 (定理 6.15-1) 是简短且具有启发性的, 但它深刻地依赖 J. L. Lions 的引理 (定理 6.11-4), 而这个引理的证明却是困难的; 另一方面, 也存在一些不依赖 J. L. Lions 引理的更直接的证明, 但它们却建筑在一些精细的计算与估计之上[64] (此类的一个证明在习题 6.15-4 中给出).

在下面, 向量值场及对称矩阵值场的空间分别用粗体的大写字母及黑板粗体大写罗马字母表示, 而其范数仍表示如同数量情况那样. 例如,

$$\|\boldsymbol{v}\|_{1,\Omega} = \left(\sum_{i=1}^{N} \|v_i\|_{1,\Omega}^2\right)^{1/2}, \text{ 对每个 } \boldsymbol{v} = (v_i) \in \boldsymbol{H}^1(\Omega) := H^1(\Omega; \mathbb{R}^N),$$

$$\|\boldsymbol{e}\|_{0,\Omega} = \left(\sum_{i,j=1}^{N} \|e_{ij}\|_{0,\Omega}^2\right)^{1/2}, \text{ 对每个 } \boldsymbol{e} = (e_{ij}) \in \mathbb{L}^2(\Omega) := L^2(\Omega; \mathbb{S}^N),$$

其中 $\mathbb{S}^N$ 表示所有实 $N \times N$ 对称矩阵的空间.

定理 6.15-1 (Korn 不等式, 又称 $\boldsymbol{H}^1(\Omega)$ 中的 Korn 不等式) *设 Ω 是 $\mathbb{R}^N$ 中的区域*[65]*. 则存在常数 $C = C(\Omega)$ 使得*

$$\|\boldsymbol{v}\|_{1,\Omega} \leqslant C(\|\boldsymbol{v}\|_{0,\Omega}^2 + \|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}^2)^{1/2}, \quad \text{对所有 } \boldsymbol{v} \in \boldsymbol{H}^1(\Omega),$$

其中

$$\boldsymbol{e}(\boldsymbol{v}) := (e_{ij}(v)) \quad \text{而} \quad e_{ij}(\boldsymbol{v}) := \frac{1}{2}(\partial_j v_i + \partial_i v_j), \quad 1 \leqslant i, j \leqslant N.$$

证明 [66] (i) *定义空间*

$$\boldsymbol{E}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega); \boldsymbol{e}(\boldsymbol{v}) \in \mathbb{L}^2(\Omega)\}.$$

则装备以范数

$$\|\boldsymbol{v}\| := (\|\boldsymbol{v}\|_{0,\Omega}^2 + \|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}^2)^{1/2}, \quad \text{对每个 } \boldsymbol{v} \in \boldsymbol{E}(\Omega),$$

空间 $\boldsymbol{E}(\Omega)$ 是 Hilbert 空间.

出现在空间 $\boldsymbol{E}(\Omega)$ 定义中的关系式 $\boldsymbol{e}(\boldsymbol{v}) \in \mathbb{L}^2(\Omega)$ 理解为在分布意义下, 即意味着存在空间 $L^2(\Omega)$ 中的函数, 将其以 $e_{ij}(\boldsymbol{v}) = e_{ji}(\boldsymbol{v})$ 表示, 使得

$$\int_\Omega e_{ij}(\boldsymbol{v})\varphi dx = -\frac{1}{2}\int_\Omega (v_i\partial_j\varphi + v_j\partial_i\varphi)dx, \quad \text{对所有 } \varphi \in \mathcal{D}(n).$$

[64] 例如, 可见 FICHERA [1972a] 或

J. A. NITSCHE [1981]: On Korn's second inequality. RAIRO Analyse Numérique **15**, 237–248.

对于具有 $\mathcal{C}^1$ 类边界的区域, 关于 J. A. Nitsche 的处理方法, 读者可在 CHIPOT [2002, 6.1 节] 中发现一个具有启发性的阐述.

[65] 说明如果 Ω 不是区域则 Korn 不等式不一定成立的反例, 可在下文中找到:

G. GEYMONAT; G. GILARDI [1998]: Contre-exemple à l'inégalité de Korn et au lemme de Lions dans des domaines irréguliers. Equations aux Dérivées Partielles et Applications. Articles Dédiés à Jacques-Louis Lions, pp. 541–548, Gauthier-Villars, Paris.

[66] 这里给出的证明出自 DUVAUT & Lions [1976, 第 3 章第 3 节] 中的定理 3.3.

考察元素 $\boldsymbol{v}^k=(v_i^k)_{i=1}^N\in\boldsymbol{E}(\Omega)$ 组成的 Cauchy 序列 $(\boldsymbol{v}^k)_{k=1}^\infty$. 范数 $\|\cdot\|$ 的定义说明, 存在函数 $v_i\in L^2(\Omega)$ 及 $e_{ij}\in L^2(\Omega)$ 使得当 $k\to\infty$ 时,

$$v_i^k\to v_i \text{ 在 } L^2(\Omega) \text{ 中} \quad \text{及} \quad e_{ij}(\boldsymbol{v}^k)\to e_{ij} \text{ 在 } L^2(\Omega) \text{ 中},$$

此因空间 $L^2(\Omega)$ 是完备的. 给定一函数 $\varphi\in\mathcal{D}(\Omega)$, 在关系式

$$\int_\Omega e_{ij}(\boldsymbol{v}^k)\varphi dx=-\frac{1}{2}\int_\Omega(v_i^k\partial_j\varphi+v_j^k\partial_i\varphi)dx, \quad k\geqslant 1,$$

中令 $k\to\infty$, 即证得 $e_{ij}=e_{ij}(\boldsymbol{v})$.

(ii) 两个空间 $\boldsymbol{E}(\Omega)$ 与 $\boldsymbol{H}^1(\Omega)$ 等同.

显然 $\boldsymbol{H}^1(\Omega)\subset\boldsymbol{E}(\Omega)$. 为了证明另一方向的包含, 设 $\boldsymbol{v}=(v_i)_{i=1}^N\in\boldsymbol{E}(\Omega)$. 则对于 $1\leqslant i,j,k\leqslant N$ 有

$$\partial_k v_i\in H^{-1}(\Omega),$$
$$\partial_j(\partial_k v_i)=\{\partial_j e_{ik}(\boldsymbol{v})+\partial_k e_{ij}(\boldsymbol{v})-\partial_i e_{jk}(\boldsymbol{v})\}\in H^{-1}(\Omega),$$

此因 $w\in L^2(\Omega)$ 意味着 $\partial_\ell w\in H^{-1}(\Omega), 1\leqslant\ell\leqslant N$. 所以由 J. L. Lions 引理 (定理 6.11-4), $\partial_k v_i\in L^2(\Omega)$, 故 $\boldsymbol{v}\in\boldsymbol{H}^1(\Omega)$.

(iii) Korn 不等式.

从装备以 $\|\cdot\|_{1,\Omega}$ 的空间 $\boldsymbol{H}^1(\Omega)$ 到装备以 $\|\cdot\|$ 的空间 $\boldsymbol{E}(\Omega)$ 的典范内射 ι 是连续的 (显然存在常数 c 使得 $\|\boldsymbol{v}\|\leqslant c\|\boldsymbol{v}\|_{1,\Omega}$ 对所有 $\boldsymbol{v}\in\boldsymbol{H}^1(\Omega)$) 单射, 且根据 (ii) 还是满射.

因此, Banach 开映射定理的推论 (定理 5.6-2) 说明, 其逆映射 ι^{-1} 也是连续的. 而这正是 Korn 不等式所描述的结果. □

在 $\mathbb{R}^N$ 中的区域 Ω 上, 还可以确立类似的不等式. 例如, $\boldsymbol{W}^{1,p}(\Omega)$ 中的 Korn 不等式, 其断言, 对每个 $1<p<\infty$, 存在常数 C_p 使得[67)]

$$\|\boldsymbol{v}\|_{1,p,\Omega}\leqslant C_p(\|\boldsymbol{v}\|_{0,p,\Omega}^p+\|\boldsymbol{e}(\boldsymbol{v})\|_{0,p,\Omega}^p)^{1/p}, \quad \text{对所有 } \boldsymbol{v}\in\boldsymbol{W}^{1,p}(\Omega)$$

或 $\boldsymbol{L}^2(\Omega)$ 中的 Korn 不等式 (这个不等式将在定理 6.19-2 的证明过程中确立), 其断言, 存在常数 C 使得

$$\|\boldsymbol{v}\|_{0,\Omega}\leqslant C(\|\boldsymbol{v}\|_{-1,\Omega}^2+\|\boldsymbol{e}(\boldsymbol{v})\|_{-1,\Omega}^2)^{1/2}, \quad \text{对所有 } \boldsymbol{v}\in\boldsymbol{L}^2(\Omega).$$

我们下一个目标是确立 $\boldsymbol{H}^1(\Omega)$ 中 Korn 不等式的一个等价形式, 一个在商空间中的不等式 (定理 6.15-3). 为此, 我们需要对于在 $\mathbb{L}^2(\Omega)$ 中满足 $\boldsymbol{e}(\boldsymbol{v})=\boldsymbol{0}$ 的向量场 $\boldsymbol{v}\in\boldsymbol{H}^1(\Omega)$, 给出一个等同描述 (定理 6.15-2). 符号 $\mathbb{A}^N$ 表示所有 $N\times N$ 实反对称矩阵的空间.

[67)] G. Geymonat; P. Suquet [1986]: Functional spaces for Norton-Hoff materials. Mathematical Methods in the Applied Sciences, **8**, 206–222.

定理 6.15-2 设 Ω 是 $\mathbb{R}^N$ 中的连通开子集. 则

$$\begin{aligned}&\{\boldsymbol{v}\in \boldsymbol{H}^1(\Omega); \boldsymbol{e}(\boldsymbol{v})=\mathbf{0} \text{ 在 } \Omega \text{ 中}\}\\ =&\{\boldsymbol{v}\in \boldsymbol{H}^1(\Omega); \text{ 存在 } \boldsymbol{B}\in\mathbb{A}^N \text{ 及 } \boldsymbol{c}\in\mathbb{R}^N \text{ 使得}\\ &\boldsymbol{v}(x)=\boldsymbol{B}x+\boldsymbol{c} \text{ 对几乎所有 } x\in\Omega\}.\end{aligned}$$

证明 对每组 $1\leqslant i,j,k\leqslant N$, 任一个向量场 $\boldsymbol{v}=(v_i)\in \boldsymbol{H}^1(\Omega)$ 都满足

$$\int_\Omega(\partial_j v_i)\partial_k\varphi dx=\int_\Omega\{e_{ij}(\boldsymbol{v})\partial_k\varphi+e_{ik}(\boldsymbol{v})\partial_j\varphi-e_{jk}(\boldsymbol{v})\partial_i\varphi\}dx \text{ 对所有 } \varphi\in\mathcal{D}(\Omega),$$

这是因为上式两端都等于 $-\int_\Omega v_i\partial_{kj}\varphi dx$ (为验证这个结论, 只需简单地利用函数 $e_{ij}(\boldsymbol{v})$ 的定义, 弱导数的定义, 并注意到如果 $\varphi\in\mathcal{D}(\Omega)$, 则每个函数 $\partial_k\varphi$ 都属于 $\mathcal{D}(\Omega)$). 所以

$$\boldsymbol{e}(\boldsymbol{v})=\mathbf{0} \text{ 在 } \Omega \text{ 中意味着 } \int_\Omega(\partial_j v_i)\partial_k\varphi dx=0, \text{ 对所有 } \varphi\in\mathcal{D}(\Omega).$$

因为 $\partial_j v_i\in L^2(\Omega)\subset L^1_{\mathrm{loc}}(\Omega)$, 故存在常数 $b_{ij}, 1\leqslant i,j\leqslant N$, 使得

$$\partial_j v_i(x)=b_{ij}, \text{ 对几乎所有 } x\in\Omega$$

(根据定理 6.3-4; 注意区域是连通的). 另外, $e_{ij}(\boldsymbol{v})=0$ 意味着

$$b_{ij}=-b_{ji}.$$

设

$$w_i(x):=\sum_{j=1}^N b_{ij}x_j, \text{ 对每个 } x\in\Omega,\quad 1\leqslant i\leqslant N.$$

则有

$$\begin{aligned}\int_\Omega v_i\partial_j\varphi dx&=-\int_\Omega(\partial_j v_i)\varphi dx=-b_{ij}\int_\Omega\varphi dx\\&=-\int_\Omega(\partial_j w_i)\varphi dx=\int_\Omega w_i\partial_j\varphi dx\end{aligned}$$

对所有 $\varphi\in\mathcal{D}(\Omega)$ (根据弱导数 $\partial_j v_i$ 的定义). 故存在常数 c_i 使得 $(v_i-w_i)(x)=c_i, 1\leqslant i\leqslant N$, 对几乎所有 $x\in\Omega$ 成立 (仍由定理 6.3-4).

所以, 我们已经证明了, 如果一个向量场 $\boldsymbol{v}\in\boldsymbol{H}^1(\Omega)$ 满足 $\boldsymbol{e}(\boldsymbol{v})=\mathbf{0}$ 在 Ω 中, 一定存在反对称的 $N\times N$ 矩阵 $\boldsymbol{B}=(b_{ij})$ 及一个向量 $\boldsymbol{c}\in\mathbb{R}^N$ 使得

$$\boldsymbol{v}(x)=\boldsymbol{B}x+\boldsymbol{c}, \quad \text{对几乎所有 } x\in\Omega.$$ □

注 如果用弱导数来表示, 在上述证明中的第一个关系式说明, 对每组 $1\leqslant i,j,k\leqslant N$, 有

$$\partial_{jk}v_i=\partial_k e_{ij}(\boldsymbol{v})+\partial_j e_{ik}(\boldsymbol{v})-\partial_i e_{jk}(\boldsymbol{v}) \quad \text{在 } H^{-1}(\Omega) \text{ 中},$$

因此它也在分布意义下成立.

当 $N=3$ 时, 定理 6.15-2 蕴涵着, 一个向量场 $\boldsymbol{v}\in\boldsymbol{H}^1(\Omega)$ 满足 $\boldsymbol{e}(\boldsymbol{v})=\mathbf{0}$ 在 $\mathbb{L}_s^2(\Omega)$ 中的充分必要条件是存在两个向量 $\boldsymbol{b}\in\mathbb{R}^3$ 及 $\boldsymbol{c}\in\mathbb{R}^3$ 使得

$$\boldsymbol{v}(x)=\boldsymbol{b}\wedge\boldsymbol{ox}+\boldsymbol{c},\quad \text{对几乎所有 } x\in\Omega$$

(其 "充分性" 部分可直接验证). 当把集合 Ω 视为一个位移场时, 这个向量场称为*无穷小刚体位移*[68].

设 $\mathbb{M}^N$ 表示所有 $N\times N$ 实矩阵的空间. 给定 $\mathbb{R}^N$ 的一个开子集及一个足够光滑的向量场 $\boldsymbol{v}=(v_i):\Omega\to\mathbb{R}^N$, $\boldsymbol{v}$ 的*梯度矩阵场*是由 $(\boldsymbol{\nabla v})_{ij}:=\partial_j v_i$ 定义的矩阵场 $\boldsymbol{\nabla v}:\Omega\to\mathbb{M}^N$. 因此在本节中引入的矩阵场 $\boldsymbol{e}(\boldsymbol{v}):\Omega\to\mathbb{S}^N$ 也可由下式定义:

$$\boldsymbol{e}(\boldsymbol{v}):=\frac{1}{2}(\boldsymbol{\nabla v}^{\mathrm{T}}+\boldsymbol{\nabla v}).$$

由于这个原因, $\boldsymbol{e}(\boldsymbol{v})$ 也称为 $\boldsymbol{v}$ 的*对称化梯度场*, 而且也将用更 "类算子" 的符号表示 (如在下述定理中):

$$\boldsymbol{\nabla}_s\boldsymbol{v}:=\frac{1}{2}(\boldsymbol{\nabla v}^{\mathrm{T}}+\boldsymbol{\nabla v}).$$

定理 6.15-3 (在商空间 $\boldsymbol{H}^1(\Omega)/\mathbf{Ker}\,\boldsymbol{\nabla}s$ 中的 Korn 不等式) *设 Ω 是 $\mathbb{R}^N$ 中的区域. 定义商空间*

$$\boldsymbol{H}^1(\Omega):=\boldsymbol{H}^1(\Omega)/\mathbf{Ker}\,\boldsymbol{\nabla}_s,$$

其中

$$\mathbf{Ker}\,\boldsymbol{\nabla}_s:=\{\boldsymbol{v}\in\boldsymbol{H}^1(\Omega);\nabla_s\boldsymbol{v}=0 \text{ 在 } \Omega \text{ 中}\}.$$

装备以如下定义的商范数

$$\|\dot{\boldsymbol{v}}\|_{1,\Omega}:=\inf_{\boldsymbol{r}\in\mathbf{Ker}\,\boldsymbol{\nabla}_s}\|\boldsymbol{v}+\boldsymbol{r}\|_{1,\Omega},\quad \text{对每个 } \dot{\boldsymbol{v}}\in\dot{\boldsymbol{H}}^1(\Omega),$$

空间 $\dot{\boldsymbol{H}}^1(\Omega)$ 是 Hilbert 空间 (习题 4.1-5). 则

(a) 由 $\boldsymbol{H}^1(\Omega)$ 中的 Korn 不等式 (定理 6.15-1) 可得, 存在常数 $\dot{C}=\dot{C}(\Omega)$ 使得在 $\dot{\boldsymbol{H}}^1(\Omega)$ 中的 Korn 不等式成立, 即

$$\|\dot{\boldsymbol{v}}\|_{1,\Omega}\leqslant\dot{C}\|\boldsymbol{e}(\dot{\boldsymbol{v}})\|_{0,\Omega},\ \text{对所有 } \dot{\boldsymbol{v}}\in\dot{\boldsymbol{H}}^1(\Omega),$$

其中 $\boldsymbol{e}(\dot{\boldsymbol{v}}):=\boldsymbol{e}(\boldsymbol{w})$ 对任意 $\boldsymbol{w}\in\dot{\boldsymbol{v}}$.

(b) 反之, 由 $\dot{\boldsymbol{H}}^1(\Omega)$ 中的 Korn 不等式可得 $\boldsymbol{H}^1(\Omega)$ 中的 Korn 不等式.

---

[68] "无穷小" 一词在此反映如下事实, 这个向量场的空间是在集合 Ω 刚体变形之流形原点处的切空间; 见下文中的定理 4.1:

P. G. CIARLET; C. MARDARE [2003]: On rigid and infinitesimal rigid displacements in three-dimensional elasticity. Mathematical Models and Methods in Applied Sciences **13**, 1589–1598.

证明 由定理 6.15-2, 空间 $\textbf{Ker}\,\boldsymbol{\nabla}_s$ 是有限维的, 其维数是 $M := N(N+1)/2$.

由赋范向量空间中的 Hahn-Banach 定理 (定理 5.9-1), 在 $\boldsymbol{H}^1(\Omega)$ 上存在 M 个连续线性形式 $\ell_\alpha, 1 \leqslant \alpha \leqslant M$, 它们具有以下性质: 一个元素 $\boldsymbol{r} \in \textbf{Ker}\,\boldsymbol{\nabla}_s$ 等于 $\mathbf{0}$ 的充分必要条件是 $\ell_\alpha(\boldsymbol{r}) = 0, 1 \leqslant \alpha \leqslant M$. 这样, 我们就断言, 存在常数 D 使得

$$\|\boldsymbol{v}\|_{1,\Omega} \leqslant D\left(\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega} + \sum_{\alpha=1}^{M} |\ell_\alpha(\boldsymbol{v})|\right), \quad \text{对所有 } \boldsymbol{v} \in \boldsymbol{H}^1(\Omega).$$

而由这个不等式又立即可得在 $\dot{\boldsymbol{H}}^1(\Omega)$ 中的 Korn 不等式: 给定任意 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$, 令 $\boldsymbol{r}(\boldsymbol{v}) \in \textbf{Ker}\,\boldsymbol{\nabla}_s$ 并使得 $\ell_\alpha(\boldsymbol{v} + \boldsymbol{r}(\boldsymbol{v})) = 0, 1 \leqslant \alpha \leqslant M$; 则

$$\begin{aligned}\|\dot{\boldsymbol{v}}\|_{1,\Omega} = \inf_{\boldsymbol{r}\in\textbf{Ker}\,\boldsymbol{\nabla}_s} \|\boldsymbol{v} + \boldsymbol{r}\|_{1,\Omega} &\leqslant \|\boldsymbol{v} + \boldsymbol{r}(\boldsymbol{v})\|_{1,\Omega} \\ &\leqslant D\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega} = D\|\boldsymbol{e}(\dot{\boldsymbol{v}})\|_{0,\Omega}.\end{aligned}$$

现在证明这样的常数 D 的存在性. 若不然, 则存在 $\boldsymbol{v}^k \in \boldsymbol{H}^1(\Omega), k \geqslant 1$, 使得

$$\|\boldsymbol{v}^k\|_{1,\Omega} = 1, \quad \text{对所有 } k \geqslant 1$$

且

$$\left(\|\boldsymbol{e}(\boldsymbol{v}^k)\|_{0,\Omega} + \sum_{\alpha=1}^{M} |\ell_\alpha(\boldsymbol{v}^k)|\right) \xrightarrow[k\to\infty]{} 0.$$

根据 Rellich-Kondrachov 定理 (定理 6.6-3), 存在在 $\boldsymbol{L}^2(\Omega)$ 中收敛的子序列 $(\boldsymbol{v}^\ell)_{\ell=1}^\infty$. 因为序列 $(\boldsymbol{e}(\boldsymbol{v}^\ell))_{\ell=1}^\infty$ 也在 $\mathbb{L}^2(\Omega)$ 中收敛, 故子序列 $(\boldsymbol{v}^\ell)_{\ell=1}^\infty$ 关于范数 $\boldsymbol{v} \to (\|\boldsymbol{v}\|_{0,\Omega}^2 + \|\boldsymbol{e}(\boldsymbol{v})_{0,\Omega}^2\|)^{1/2}$ 是 Cauchy 序列. 所以由在 $\boldsymbol{H}^1(\Omega)$ 中的 Korn 不等式 (定理 6.15-1), 它关于范数 $\|\cdot\|_{1,\Omega}$ 也是 Cauchy 序列. 因此, 存在 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ 使得

$$\|\boldsymbol{v}^\ell - \boldsymbol{v}\|_{1,\Omega} \xrightarrow[\ell\to\infty]{} 0.$$

由于 $\boldsymbol{e}(\boldsymbol{v}) = 0$ 及 $\ell_\alpha(\boldsymbol{v}) = 0, 1 \leqslant \alpha \leqslant M$, 故 $\boldsymbol{v} = \mathbf{0}$. 这与对所有 $\ell \geqslant 1, \|\boldsymbol{v}^\ell\|_{1,\Omega} = 1$ 矛盾. (a) 的结论得证[69].

下面证明反之的结论, 即由商空间 $\dot{\boldsymbol{H}}^1(\Omega)$ 中的 Korn 不等式可推得空间 $\boldsymbol{H}^1(\Omega)$ 中的 Korn 不等式.

假定这个结论不成立, 则存在 $\boldsymbol{v}^k \in \boldsymbol{H}^1(\Omega), k \geqslant 1$, 使得

$$\|\boldsymbol{v}^k\|_{1,\Omega} = 1, \quad \text{对所有 } k \geqslant 1$$

且

$$\|\boldsymbol{v}^k\|_{0,\Omega}^2 + \|\boldsymbol{e}(\boldsymbol{v}^k)\|_{0,\Omega}^2{}^{1/2} \xrightarrow[k\to\infty]{} 0.$$

[69] (a) 的结论的另一证明可参阅 Duvaut & Lions [1976, 第 3 章, 定理 3.4].

令 $\boldsymbol{r}^k \in \mathbf{Ker}\,\boldsymbol{\nabla}_s$ 对每个 $k \geqslant 1$ 表示 $\boldsymbol{v}^k$ 关于 $\boldsymbol{H}^1(\Omega)$ 的内积在 $\mathbf{Ker}\,\boldsymbol{\nabla}_s$ 上的投影，那么它应满足:

$$\|\boldsymbol{v}^k - \boldsymbol{r}^k\|_{1,\Omega} = \inf_{\boldsymbol{r}\in\mathbf{Ker}\,\boldsymbol{\nabla}_s} \|\boldsymbol{v}^k - \boldsymbol{r}\|_{1,\Omega}$$

及

$$\|\boldsymbol{v}^k\|_{1,\Omega}^2 = \|\boldsymbol{v}^k - \boldsymbol{r}^k\|_{1,\Omega} + \|\boldsymbol{r}^k\|_{1,\Omega}^2.$$

空间 $\mathbf{Ker}\,\boldsymbol{\nabla}_s$ 是有限维的, 故不等式 $\|\boldsymbol{r}^k\|_{1,\Omega} \leqslant 1$ 对所有 $k \geqslant 1$ 成立意味着存在子序列 $(\boldsymbol{r}^\ell)_{\ell=1}^\infty$, 它在 $\boldsymbol{H}^1(\Omega)$ 中收敛于一个元素 $\boldsymbol{r} \in \mathbf{Ker}\,\boldsymbol{\nabla}_s$. 此外, 在 $\dot{\boldsymbol{H}}^1(\Omega)$ 中的 Korn 不等式意味着 $\|\boldsymbol{v}^\ell - \boldsymbol{r}^\ell\|_{1,\Omega} \xrightarrow[\ell\to\infty]{} 0$, 故有

$$\|\boldsymbol{v}^\ell - \boldsymbol{r}\|_{1,\Omega} \xrightarrow[\ell\to\infty]{} 0.$$

因此 $\|\boldsymbol{v}^\ell - \boldsymbol{r}\|_{0,\Omega} \xrightarrow[\ell\to\infty]{} 0$, 另一方面有 $\|\boldsymbol{v}^\ell\|_{0,\Omega} \to 0$, 这就得到 $\boldsymbol{r}$ 等于 $\mathbf{0}$. 这样, 我们就得到 $\|\boldsymbol{v}^\ell\|_{1,\Omega} \to 0$, 导致矛盾. □

最后, 我们考察 (齐次) 边界条件的效应.

我们已经知道, 如果 $d\Gamma$-meas-$\Gamma_0 > 0$, 在空间 $H^1(\Omega)$ 的闭子空间 $\{v \in H^1(\Omega); v = 0$ 在 Γ_0 上$\}$ 上, 半范数 $|\cdot|_{1,\Omega}$ 就成为等价于 $\|\cdot\|_{1,\Omega}$ 的范数 (定理 6.6-6). 在下述定理中将要证明, 如果 $d\Gamma$-meas-$\Gamma_0 > 0$, 在 $\boldsymbol{H}^1(\Omega)$ 的闭子空间 $\{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega); \boldsymbol{v} = \mathbf{0}$ 在 Γ_0 上$\}$ 上, 半范数 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \to \|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}$ 也类似地成为一个等价于 $\|\cdot\|_{1,\Omega}$ 的范数.

值得注意的是, 尽管对于任意子集 $\Gamma_0 \subset \Gamma$, $d\Gamma$-meas-$\Gamma_0 > 0$ 情况的证明依赖 Korn 不等式 (定理 6.15-1), 而后者的证明建立在 J. L. Lions 引理之上. 但是对于 $\Gamma_0 = \Gamma$ 这一特殊情况, 其证明变得令人难以置信地容易, 可视为一个简单恒等式的直接推论 (习题 6.15-1).

定理 6.15-4 (有边界条件的 Korn 不等式) 设 Ω 是 $\mathbb{R}^N$ 中的区域, Γ_0 是 Ω 的边界 Γ 的一个 $d\Gamma$ 可测子集且 $d\Gamma$-meas-$\Gamma_0 > 0$. 则空间

$$\boldsymbol{V} := \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega); \boldsymbol{v} = \mathbf{0} \text{ 在 } \Gamma_0 \text{ 上}\}$$

是 $\boldsymbol{H}^1(\Omega)$ 的闭子空间, 而且存在常数 $C = C(\Omega, \Gamma_0)$ 使得

$$\|\boldsymbol{v}\|_{1,\Omega} \leqslant C\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega} \quad \text{对所有 } \boldsymbol{v} \in \boldsymbol{V}.$$

证明 (i) 空间 $\boldsymbol{V}$ 在 $\boldsymbol{H}^1(\Omega)$ 中是闭的, 而且由下式定义的半范数 $|\cdot| : \boldsymbol{H}^1(\Omega) \to \mathbb{R}$,

$$|\boldsymbol{v}| := \|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}, \quad \text{对每个 } \boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$$

在空间 $\boldsymbol{V}$ 上是范数.

$\boldsymbol{V}$ 在 $\boldsymbol{H}^1(\Omega)$ 中是闭的这一论断, 可如在定理 6.7-5 的证明中那样予以论证. 我们在定理 6.15-2 中已知, 如果一向量场 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ 在 Ω 中满足 $\boldsymbol{e}(\boldsymbol{v}) = \mathbf{0}$, 则存在一个 $N \times N$ 反对称矩阵 $\boldsymbol{B}$ 及一个向量 $\boldsymbol{c} \in \mathbb{R}^N$ 使得

$$\boldsymbol{v}(x) = \boldsymbol{B}x + \boldsymbol{c}, \quad \text{对几乎所有 } x \in \Omega.$$

又因为, 除了 $\boldsymbol{B}=\mathbf{0}, \boldsymbol{c}=\mathbf{0}$ 的情况, 如果这样一个向量场 $\boldsymbol{v}$ 在 $\mathbb{R}^N$ 的一个子集上为零向量, 则该子集一定是零 $\mathbb{R}^{N-1}$ 测度的 (习题 6.15-2), 因此若 $\mathrm{d}\Gamma\text{-meas-}\Gamma_0>0$ 则有 $\boldsymbol{v}=\mathbf{0}$. 所以半范数 $|\cdot|$ 是空间 $\boldsymbol{V}$ 上的范数.

(ii) 有边界条件的 Korn 不等式.

如果这个不等式不成立, 则一定存在元素 $\boldsymbol{v}^k \in \boldsymbol{V}$ 的序列 $(\boldsymbol{v}^k)_{k=1}^{\infty}$, 使得

$$\|\boldsymbol{v}^k\|_{1,\Omega}=1, \quad 对所有\ k \geqslant 1$$

且

$$\lim_{k\to\infty}\|\boldsymbol{e}(\boldsymbol{v}^k)\|_{0,\Omega}=0.$$

序列 $(\boldsymbol{v}^k)_{k=1}^{\infty}$ 在 $\boldsymbol{H}^1(\Omega)$ 中有界, 由 Rellich-Kondrachov 定理 (定理 6.6-3) 知, 存在一个在 $\boldsymbol{L}^2(\Omega)$ 中收敛的子序列 $(\boldsymbol{v}^\ell)_{\ell=1}^{\infty}$; 而且序列 $(\boldsymbol{e}(\boldsymbol{v}^\ell))_{\ell=1}^{\infty}$ 也在 $\boldsymbol{L}^2(\Omega)$ 中收敛 (到 $\mathbf{0}$, 但在现阶段用不到这一事实). 所以子序列 $(\boldsymbol{v}^\ell)_{\ell=1}^{\infty}$ 关于如下定义的范数 $\|\cdot\|$:

$$\|\boldsymbol{v}\|:=(\|\boldsymbol{v}\|_{0,\Omega}^2+\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}^2)^{1/2}, \quad 对每个\ \boldsymbol{v}\in \boldsymbol{H}^1(\Omega)$$

是 Cauchy 序列. 根据 $\boldsymbol{H}^1(\Omega)$ 中的 Korn 不等式 (定理 6.15-1), 该子序列关于范数 $\|\cdot\|_{1,\Omega}$ 也是 Cauchy 序列.

空间 $\boldsymbol{V}$ (作为 $\boldsymbol{H}^1(\Omega)$ 的闭子空间) 是完备的, 故存在 $\boldsymbol{v}\in\boldsymbol{V}$ 使得

$$\boldsymbol{v}^\ell \to \boldsymbol{v}\ 在\ \boldsymbol{H}^1(\Omega)\ 中, \quad 当\ \ell\to\infty\ 时,$$

这样极限 $\boldsymbol{v}$ 就满足 $\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}=\lim_{\ell\to\infty}\|\boldsymbol{e}(\boldsymbol{v}^\ell)\|_{0,\Omega}=0$, 由 (i) 得 $\boldsymbol{v}=\mathbf{0}$. 但这与关系式 $\|\boldsymbol{v}^\ell\|_{1,\Omega}=1$ 对所有 $\ell\geqslant 1$ 成立矛盾. 证明完成. □

在下一节中将会看到, 上面建立的 Korn 不等式 ($N=3$ 的情况) 在三维线性化弹性存在性定理的证明中起着实质性的作用. 同样地, 我们可建立另外一些 Korn 不等式, 它们在证明线性化壳理论存在性定理中起着实质性的作用. 这其中包括: 在曲面 (8.8 节中所定义的曲面) 上的一般 Korn 不等式[70]; 椭圆曲面 (即其上的所有点都

---

[70] 应归之于:

M. BERNADOU; P. G. CIARLET [1976]: Sur l'ellipticité du modèle linéaire de coques de W. T. Koiter. Computing Methods in Applied Sciences and Engineering (R. GLOWINSKI & J. L. LIONS, editors), pp. 89–136, Lecture Notes in Economics and Mathematical Systems, **134**, Springer, Heidelberg.

其他的证明及推广, 可见下述论文:

M. BERNADOU; P. G. CIARLET; B. MIARA [1994]: Existence theorems for two-dimensional linear shell theories. Journal of Elasticity **34**, 111–138.

A. BLOUZA; H. LE DRET [1999]: Existence and uniqueness for the linear Koiter model for shells with little regularity. Quarterly of Applied Mathematics **57**, 317–337.

P. G. CIARLET; S. MARDARE [2001]: On Korn's inequalities in curvilinear coordinates. Mathematical Models and Methods in Applied Sciences **11**, 1379–1391.

J. L. AKIAN [2003]; A simple proof of the ellipticity of Koiter's model. Analysis and Applications **1**, 1–16.

是“椭圆的”; 见 8.12 节) 上的 Korn 不等式[71]; 在无边界曲面上的 Korn 不等式[72]; 在无边界的椭圆曲面上的 Korn 不等式[73]; 或 Riemann 流形上的 Korn 不等式[74].

习题

6.15-1 设 Ω 是 $\mathbb{R}^N$ 中的开子集.

(1) 给定一个向量场 $\boldsymbol{v}=(v_i)_{i=1}^N:\Omega\to\mathbb{R}^N$, 其分量 $v_i\in\mathcal{C}^\infty(\Omega)$, 证明在 Ω 中成立

$$2\sum_{i,j=1}^N|e_{ij}(\boldsymbol{v})|^2-\sum_{i,j=1}^N|\partial_iv_j|^2=\left|\sum_{i=1}^N\partial_iv_i\right|^2+\sum_{i,j=1}^N\partial_i(v_j\partial_jv_i-v_i\partial_jv_j).$$

(2) 验证由 (1) 可推得

$$|\boldsymbol{v}|_{1,\Omega}\leqslant\sqrt{2}\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega},\quad 对所有\ \boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega).$$

(3) 如果 Ω 具有有限宽度 (6.5 节), 证明半范数 $\boldsymbol{v}\in\boldsymbol{H}_0^1(\Omega)\to\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}$ 就成为空间 $\boldsymbol{H}_0^1(\Omega)$ 上的范数, 且其等价于范数 $\|\cdot\|_{1,\Omega}$ (这个结果构成定理 6.15-4 中 $\Gamma_0=\Gamma$ 的一个特殊情况, 但在此关于 Ω 的假设较弱, 因定理中假设 Ω 是一个区域).

6.15-2 (1) 给定两个向量 $\boldsymbol{b},\boldsymbol{c}\in\mathbb{R}^3$. 证明, 除了 $\boldsymbol{b}=\boldsymbol{c}=\boldsymbol{0}$ 的情况, 集合 $\boldsymbol{E}:=\{x\in\mathbb{R}^3;\boldsymbol{b}\wedge\boldsymbol{ox}+\boldsymbol{c}=\boldsymbol{0}\}$ 的面积为零.

提示: 证明如果 $\boldsymbol{b}=\boldsymbol{0}$ 但 $\boldsymbol{c}\neq\boldsymbol{0}$, 或 $\boldsymbol{b}\neq\boldsymbol{0}$ 且 $\boldsymbol{b}\cdot\boldsymbol{c}\neq0$, 则 $\boldsymbol{E}=\varnothing$. 然后证明若 $\boldsymbol{b}\neq\boldsymbol{0}$ 但 $\boldsymbol{b}\cdot\boldsymbol{c}=0$, 则 $\boldsymbol{E}=\left\{\left(\dfrac{\boldsymbol{b}\wedge\boldsymbol{c}}{\boldsymbol{b}\cdot\boldsymbol{b}}+t\boldsymbol{b}\right)\in\mathbb{R}^3;t\in\mathbb{R}\right\}$.

(2) 更一般地, 证明: 给定 $N\times N$ 反对称矩阵 $\boldsymbol{B}$ 和向量 $\boldsymbol{c}\in\mathbb{R}^N$, 集合 $\boldsymbol{E}:=\{\boldsymbol{x}\in\mathbb{R}^N;\boldsymbol{Bx}+\boldsymbol{c}=\boldsymbol{0}\}$ 是零 $\mathbb{R}^{N-1}$ Lebesgue 测度的, 除非 $\boldsymbol{B}=\boldsymbol{0}$ 且 $\boldsymbol{c}=\boldsymbol{0}$ (这个结果在定理 6.15-4 的证明 (i) 部分中用到过).

6.15-3 设 ω 是 $\mathbb{R}^2$ 中的区域. 定理 6.15-1 证明了存在常数 $c=c(\omega)>0$ 使得

$$\|\boldsymbol{\eta}\|_{1,\omega}\leqslant c(\|\boldsymbol{\eta}\|_{0,\omega}^2+\|\boldsymbol{e}(\boldsymbol{\eta})\|_{0,\omega}^2)^{1/2},\quad 对所有\ \boldsymbol{\eta}=(\eta_\alpha)\in H^1(\omega)\times H^1(\omega),$$

其中 $e(\boldsymbol{\eta})=(e_{\alpha\beta}(\boldsymbol{\eta})),e_{\alpha\beta}(\boldsymbol{\eta}):=\dfrac{1}{2}(\partial_\alpha\eta_\beta+\partial_\beta\eta_\alpha)$. 证明, 这个 $\boldsymbol{H}^1(\omega)$ 中的二维 Korn 不等式也可从在 $\boldsymbol{H}^1(\omega\times]-1,1[\ )$ 中的三维 Korn 不等式导出.

---

[71] P. G. Ciarlet; V. Lods [1996]: On the ellipticity of linear membrane shell equations. Journal de Mathématiques Pures et Appliquées **75**, 107–124.

P. G. Ciarlet; E. Sanchez-Palencia [1996]: An existence and uniqueness theorem for the two-dimensional linear membrane shell equation. Journal de Mathématiques Pures et Appliquées **75**, 51–67.

[72] S. Mardare [2003]: Inequality of Korn's type on compact surfaces without boundary. Chinese Annals of Mathematics, Series B, **24**, 191–204.

[73] S. Slicaru [1998]: On the ellipticity of the middle surface of a shell and its application to the asymptotic analysis of “membrane shells”. Journal of Elasticity **46**, 33–42.

[74] W. Chen; J. Jost [2002]; A Riemannian version of Korn's inequality. Calculus of Variations **14**, 517–530.

6.15-4 在 1988 年, V. A. Kondrat'ev 与 O. Oleinik 发表了关于 Korn 不等式的一个值得关注的证明[75], 这个证明是自包含的, 而且在很大程度上是初等的. 他们的证明正是本习题要讨论的目标, 它不依赖高等泛函分析的结果, 譬如 (定理 6.15-1 中所用的) J. L. Lions 引理, 或在 1962 年 J. Gobert 给出的最初证明 (本节初曾引用过) 中所用的奇异积分 Calderón-Zygmund 理论[76], 而依赖两个关键的不等式及 Laplace 算子 Δ 的次椭圆性 (定理 6.4-2). 这两个不等式构成下面的问题 (1) 及 (2) (然而, 其证明, 尤其是第二个不等式的证明有些难度), 而在问题 (2) 中要用到次椭圆性.

在下面 Ω 是 $\mathbb{R}^N$ 中的区域, 符号 C_1, C_2 等表示只依赖 Ω 的各种常数, 函数 $\rho : \overline{\Omega} \to \mathbb{R}$ 定义为 $\rho(x) := \operatorname{dist}(x, \partial\Omega)$ 对每个 $x \in \overline{\Omega}$, 而 $\boldsymbol{u} = (u_i)_{i=1}^N \in \boldsymbol{C}^\infty(\overline{\Omega})$ 表示一个给定的向量场. 其他的符号与正文中相同.

(1) 证明

$$\int_\Omega \rho^2 \sum_{k=1}^N |\partial_k u|^2 dx \leqslant C_1(\|v\|_{0,\Omega}^2 + \|\Delta v\|_{0,\Omega}^2)$$

对所有满足 $\Delta v \in L^2(\Omega)$ 的函数 $v \in L^2(\Omega) \cap \mathcal{C}^\infty(\Omega)$ 成立 (对所有这种函数 v, 这个不等式的右端显然是有限的).

(2) 证明

$$|v|_{1,\Omega}^2 \leqslant C_2 \left(\int_\Omega \rho^2 \sum_{i,j=1}^N |\partial_{ij} v|^2 dx + \|v\|_{0,\Omega}^2 \right)$$

对所有满足 $\int_\Omega \rho^2 \sum_{i,j=1}^N |\partial_{ij} v|^2 dx < \infty$ 的函数 $v \in H^1(\Omega) \bigcap \mathcal{C}^\infty(\Omega)$ 成立.

(3) 构造一个向量场 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \bigcap \boldsymbol{C}^\infty(\Omega)$ 使其满足

$$\Delta \boldsymbol{v} = \Delta \boldsymbol{u} \quad \text{在 } \Omega \text{ 中} \quad \text{且 } \|\boldsymbol{v}\|_{1,\Omega} \leqslant C_3 \|\boldsymbol{e}(\boldsymbol{u})\|_{0,\Omega}.$$

提示: 利用关系式 $\Delta u_i = \sum_{j=1}^N (2\partial_j e_{ij}(\boldsymbol{u}) - \partial_i e_{jj}(\boldsymbol{u})), 1 \leqslant i \leqslant N$.

(4) 设 $\boldsymbol{w} := \boldsymbol{u} - \boldsymbol{v}$. 利用问题 (1) 证明, 对所有 $1 \leqslant i, j \leqslant N$, 成立

$$\int_\Omega \rho^2 \sum_{k=1}^N |\partial_k e_{ij}(\boldsymbol{w})|^2 dx \leqslant C_4 \|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}^2 \leqslant C_5 \|\boldsymbol{e}(\boldsymbol{u})\|_{0,\Omega}^2.$$

(5) 利用问题 (4) 及恒等式

$$\partial_{jk} w_i = \partial_j e_{ik}(\boldsymbol{w}) + \partial_k e_{ij}(\boldsymbol{w}) - \partial_i e_{ik}(\boldsymbol{w}) \quad \text{对所有 } 1 \leqslant i, j, k \leqslant N$$

(绝不是巧合, 同一恒等式在定理 6.15-1 的证明中也用过) 证明, 对所有 $1 \leqslant k \leqslant N$, 成立

$$\int_\Omega \rho^2 \sum_{i,j=1}^N |\partial_{ij} w_k|^2 dx \leqslant C_6 \|\boldsymbol{e}(\boldsymbol{u})\|_{0,\Omega}^2.$$

---

[75] V. A. KONDRAT'EV; O. A. OLEINIK [1988]: Boundary-value problems for the system of elasticity theory in unbouded domains. Korn's inequalities, Uspehi Mathematiceskii Nauk **43**, 55–98 (俄文) [英译文: Russian Mathematical Surveys **43** (1988), 65–119].

[76] Korn 不等式的另一证明依赖奇异积分的 Calderón-Zygmund 理论 (以及 Cesàro-Volterra 路径积分公式; 见定理 6.18-2), 属于:

P. P. MOSOLOV; V. P. MJASNIKOV [1971]: A proof of Korn's inequality. Soviet Mathematics Doklady **12**, 1618–1622.

(6) 利用问题 (2) 和 (5) 以及关系式 $\boldsymbol{w} = \boldsymbol{u} - \boldsymbol{v}$, 最后得

$$|\boldsymbol{u}|_{1,\Omega}^2 \leqslant C_7(\|\boldsymbol{u}\|_{0,\Omega}^2 + \|\boldsymbol{e}(\boldsymbol{u})\|_{0,\Omega}^2).$$

因为空间 $\mathcal{C}^\infty(\overline{\Omega})$ 在空间 $H^1(\Omega)$ 中稠密 (定理 6.6-4), 由此不等式立得 Korn 不等式.

6.16 Korn 不等式的应用: 三维线性化弹性方程组

本节的目标是建立三维线性化弹性方程组弱解的存在性定理, 这一结果将在下述定理中以通常的方式 (即确立函数空间、双线性形式以及线性形式) 给出. 为达此目标, 关键的步骤还是验证此时建立的双线性形式是强制的, 而为证明这一点, 则需要 Korn 不等式.

在本节中, 拉丁下标的取值范围是集合 $\{1,2,3\}$, 除非用于特别指明的序列. 而关于相同指标的求和约定, 也援此例.

给定一个定义在 $\overline{\Omega}$ 上的充分光滑的 3×3 矩阵场 $\boldsymbol{\sigma} = (\sigma_{ij})$, 其散度 $\mathbf{div}\,\boldsymbol{\sigma} : \overline{\Omega} \to \mathbb{R}^3$ 是由下式定义的向量场:

$$\mathbf{div}\,\boldsymbol{\sigma} := \begin{pmatrix} \partial_1\sigma_{11} + \partial_2\sigma_{12} + \partial_3\sigma_{13} \\ \partial_1\sigma_{21} + \partial_2\sigma_{22} + \partial_3\sigma_{23} \\ \partial_1\sigma_{31} + \partial_2\sigma_{32} + \partial_3\sigma_{33} \end{pmatrix}.$$

本节所用向量和矩阵场空间的符号与 6.15 节中的相同. 而矩阵内积的表示可参见 4.2 节.

定理 6.16-1 (三维线性化弹性方程组解的存在性) 设 Ω 是 $\mathbb{R}^3$ 中的区域, Γ_1 是 $\Gamma := \partial\Omega$ 的相对开子集, 使得

$$d\Gamma\text{-meas-}\Gamma_0 > 0, \quad \text{其中} \quad \Gamma_0 := \Gamma - \Gamma_1.$$

又设 λ 和 μ 是两个常数, 它们满足

$$\lambda \geqslant 0 \quad \text{及} \quad \mu > 0,$$

而

$$\boldsymbol{f} = (f_i) \in \boldsymbol{L}^2(\Omega) \quad \text{及} \quad \boldsymbol{g} = (g_i) \in \boldsymbol{L}^2(\Gamma_1)$$

是两个给定的向量场. 最后, 设

$$\boldsymbol{V} := \{\boldsymbol{v} = (v_i) \in \boldsymbol{H}^1(\Omega); \boldsymbol{v} = \boldsymbol{0} \text{ 在 } \Gamma_0 \text{ 上}\},$$
$$a(\boldsymbol{u},\boldsymbol{v}) := \int_\Omega \{\lambda \mathrm{tr}\,\boldsymbol{e}(\boldsymbol{u})\mathrm{tr}\,\boldsymbol{e}(\boldsymbol{v}) + 2\mu\boldsymbol{e}(\boldsymbol{u}) : \boldsymbol{e}(\boldsymbol{v})\}dx \quad \text{对每个 } \boldsymbol{u},\boldsymbol{v} \in \boldsymbol{V},$$

其中

$$
e(\boldsymbol{v}):=(e_{ij}(\boldsymbol{v}))\in\mathbb{L}^2(\Omega),\ \text{而}\ e_{ij}(\boldsymbol{v}):=\frac{1}{2}(\partial_j v_i+\partial_i v_j)\quad\text{对每个}\ \boldsymbol{v}=(v_i)\in\boldsymbol{H}^1(\Omega),
$$

$$
\ell(\boldsymbol{v}):=\int_\Omega \boldsymbol{f}\cdot\boldsymbol{v}dx+\int_{\Gamma_1}\boldsymbol{g}\cdot\boldsymbol{v}d\Gamma\quad\text{对所有}\ \boldsymbol{v}\in\boldsymbol{V}.
$$

则存在唯一的向量场 $\boldsymbol{u}=(u_i)\in\boldsymbol{V}$ 使如下定义的泛函 $J:\boldsymbol{V}\to\mathbb{R}$ 达到极小:

$$
\begin{aligned}
J(\boldsymbol{v}):&=\frac{1}{2}a(\boldsymbol{v},\boldsymbol{v})-\ell(\boldsymbol{v})\\
&=\frac{1}{2}\int_\Omega\{\lambda(\operatorname{tr}\boldsymbol{e}(\boldsymbol{v}))^2+2\mu\boldsymbol{e}(\boldsymbol{v}):\boldsymbol{e}(\boldsymbol{v})\}dx-\left(\int_\Omega\boldsymbol{f}\cdot\boldsymbol{v}dx+\int_{\Gamma_1}\boldsymbol{g}\cdot\boldsymbol{v}d\Gamma\right)
\end{aligned}
$$

对所有 $\boldsymbol{v}\in\boldsymbol{V}$, 或等价地满足下述变分方程

$$
\begin{aligned}
&\int_\Omega\{\lambda\operatorname{tr}\boldsymbol{e}(\boldsymbol{u})\operatorname{tr}\boldsymbol{e}(\boldsymbol{v})+2\mu\boldsymbol{e}(\boldsymbol{u}):\boldsymbol{e}(\boldsymbol{v})\}dx\\
=&\int_\Omega\boldsymbol{f}\cdot\boldsymbol{v}dx+\int_{\Gamma_1}\boldsymbol{g}\cdot\boldsymbol{v}d\Gamma\quad\text{对所有}\ \boldsymbol{v}\in\boldsymbol{V}.
\end{aligned}
$$

进而假定 $\boldsymbol{u}\in\boldsymbol{H}^2(\Omega)$, 则 $\boldsymbol{u}$ 满足边值问题:

$$
\begin{aligned}
-\operatorname{div}\{\lambda(\operatorname{tr}\boldsymbol{e}(\boldsymbol{u}))\boldsymbol{I}+2u\boldsymbol{e}(\boldsymbol{u})\}=\boldsymbol{f}&\quad\text{在}\ \Omega\ \text{内},\\
\boldsymbol{u}=\boldsymbol{0}&\quad\text{在}\ \Gamma_0\ \text{上},\\
\{\lambda(\operatorname{tr}\boldsymbol{e}(\boldsymbol{u}))\boldsymbol{I}+2u\boldsymbol{e}(\boldsymbol{u})\}\boldsymbol{\nu}=\boldsymbol{g}&\quad\text{在}\ \Gamma_1\ \text{上},
\end{aligned}
$$

其中 $\boldsymbol{\nu}=(\nu_i):\Gamma\to\mathbb{R}^3$ 表示 Γ 的单位外法向量场.

证明 空间 $\boldsymbol{V}$ 作为 $\boldsymbol{H}^1(\Omega)$ 的闭子空间 (定理 6.15-4) 是 Hilbert 空间. 利用 Cauchy-Schwarz 不等式, 对称双线性形式 $a(\cdot,\cdot)$ 及线性形式 ℓ 在空间 $\boldsymbol{H}^1(\Omega)$ 上是连续的. 而且双线性形式是 V 强制的, 此因

$$
\begin{aligned}
a(\boldsymbol{v},\boldsymbol{v})&=\int_\Omega\{\lambda(\operatorname{tr}\boldsymbol{e}(\boldsymbol{v}))^2+2\mu\boldsymbol{e}(\boldsymbol{v}):\boldsymbol{e}(\boldsymbol{v})\}dx\\
&\geqslant 2\mu\int_\Omega\boldsymbol{e}(\boldsymbol{v}):\boldsymbol{e}(\boldsymbol{v})dx=2\mu\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}^2\quad\text{对所有}\ \boldsymbol{v}\in\boldsymbol{V},
\end{aligned}
$$

又根据具有边界条件的 Korn 不等式 (定理 6.15-4), 存在常数 $C>0$ 使得 $\|\boldsymbol{e}(\boldsymbol{v})\|_{0,\Omega}\geqslant C^{-1}\|v\|_{1,\Omega}$ 对所有 $\boldsymbol{v}\in\boldsymbol{V}$.

因此, 根据定理 6.1-1 及定理 6.1-2, 存在唯一的向量场 $\boldsymbol{u}$ 使所给出的泛函 J 在其上达到在空间 $\boldsymbol{V}$ 上的极小, 或等价地满足所示的变分方程.

为了导出相应的边值问题, 对任意的 $\boldsymbol{u}=(u_i),\boldsymbol{v}=(v_i)\in\boldsymbol{H}^1(\Omega)$, 我们首先将 $a(\boldsymbol{u},\boldsymbol{v})$ 改写为如下形式:

$$
a(\boldsymbol{u},\boldsymbol{v})=\int_\Omega\sigma_{ij}(\boldsymbol{u})e_{ij}(\boldsymbol{v})dx=\int_\Omega\sigma_{ij}(\boldsymbol{u})\partial_j v_i dx,
$$

其中

$$\sigma_{ij}(\boldsymbol{u}) := \lambda \mathrm{tr}\, \boldsymbol{e}(\boldsymbol{u})\delta_{ij} + 2\mu e_{ij}(\boldsymbol{u}) = \sigma_{ji}(\boldsymbol{u}).$$

利用基本 Green 公式 (定理 6.6-7), 可知下述 Green 公式成立:

$$\int_\Omega \sigma_{ij}(\boldsymbol{u})\partial_j v_i dx = -\int_\Omega (\partial_j \sigma_{ij}(\boldsymbol{u}))v_i dx + \int_\Gamma \sigma_{ij}(\boldsymbol{u})\nu_j v_i d\Gamma \quad \text{对所有 } \boldsymbol{u} \in \boldsymbol{H}^2(\Omega), \boldsymbol{v} \in \boldsymbol{H}^1(\Omega).$$

如果 $\boldsymbol{u} \in \boldsymbol{H}^2(\Omega) \bigcap \boldsymbol{V}$, 这样, 变分方程 $a(\boldsymbol{u}, \boldsymbol{v}) = \ell(\boldsymbol{v})$ 对所有 $\boldsymbol{v} \in \boldsymbol{V}$ 就变成

$$\int_\Omega (-\partial_j \sigma_{ij}(\boldsymbol{u}) - f_i)v_i dx = \int_{\Gamma_1} (g_i - \sigma_{ij}(\boldsymbol{u})\nu_j)v_i d\Gamma \quad \text{对所有 } (v_i) \in \boldsymbol{V}.$$

特别地, 对每个 $1 \leqslant i \leqslant 3$, 有

$$\int_\Omega (-\partial_j \sigma_{ij}(\boldsymbol{u}) - f_i)v_i dx = 0 \quad \text{对所有 } v_i \in \mathcal{D}(\Omega),$$

这意味着 $-\partial_j \sigma_{ij}(\boldsymbol{u}) - f_i = 0$ 在 $L^2(\Omega)$ 中 (定理 6.3-2), 或等价地写成向量形式

$$-\mathbf{div}\, \boldsymbol{\sigma}(\boldsymbol{u}) = \boldsymbol{f} \text{ 在 } \boldsymbol{L}^2(\Omega) \text{ 中, 其中 } \boldsymbol{\sigma}(\boldsymbol{u}) := (\sigma_{ij}(\boldsymbol{u})) = \lambda(\mathrm{tr}\, \boldsymbol{e}(\boldsymbol{u}))\boldsymbol{I} + 2u\boldsymbol{e}(\boldsymbol{u}).$$

注意到这些方程, 对每个 $1 \leqslant i \leqslant 3$, 剩下来的就是

$$\int_{\Gamma_1} (g_i - \sigma_{ij}(\boldsymbol{u})\nu_j)v_i d\Gamma = 0 \quad \text{对所有 } v_i \in \{\omega \in \boldsymbol{H}^1(\Omega); w = 0 \text{ 在 } \Gamma_0 \text{ 上}\},$$

这就意味着 $g_i - \sigma_{ij}(\boldsymbol{u})\nu_j = 0$ 在 $L^2(\Gamma_1)$ 内 (定理 6.7-3); 或等价地, 其向量形式为

$$\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{\nu} = \boldsymbol{g} \quad \text{在 } \boldsymbol{L}^2(\Gamma_1) \text{ 中}.$$

最后, 由于 $\boldsymbol{u} \in \boldsymbol{V}$, 故 $\boldsymbol{u} = \mathbf{0}$ 在 Γ_0 上. □

注 (1) 如果 Lamé 常数只满足较弱的假设 $3\lambda + 2u > 0$ 及 $\mu > 0$, 双线性形式 $a(\cdot,\cdot)$ 仍保持 V 强制性; 见习题 6.16-1(1).

(2) 在 $\boldsymbol{f} \in \boldsymbol{L}^{6/5}(\Omega)$ 及 $\boldsymbol{g} \in \boldsymbol{L}^{4/3}(\Gamma_1)$ 的较弱假设下, 利用 Sobolev 嵌入定理 (定理 6.6-1) 在 $N = 3$ 的特殊情况下, 并结合逆算子的连续性 (定理 6.6-5) 可以证明, 线性形式 ℓ 保持在空间 $\boldsymbol{H}^1(\Omega)$ 上的连续性.

(3) 向量方程 $-\mathbf{div}\{\lambda(\mathrm{tr}\, \boldsymbol{e}(\boldsymbol{u})\boldsymbol{I} + 2\mu \boldsymbol{e}(\boldsymbol{u}))\} = \boldsymbol{f}$ 在 Ω 中可以等价地写为 Navier 方程[77] 的形式, 即

$$-\mu\Delta\boldsymbol{u} - (\lambda + \mu)\mathbf{grad}\, \mathrm{div}\, \boldsymbol{u} = \boldsymbol{f} \quad \text{在 } \Omega \text{ 内},$$

或

$$\mu\mathbf{curlcurl}\, \boldsymbol{u} - (\lambda + 2\mu)\mathbf{grad}\, \mathrm{div}\, \boldsymbol{u} = \boldsymbol{f} \quad \text{在 } \Omega \text{ 内}.$$

[77] 冠名源自 Claude Louis Navier (1785—1836).

在定理 6.16-1 中讨论的边值问题, 即

$$-\mathbf{div}\,\boldsymbol{\sigma}(\boldsymbol{u}) = \boldsymbol{f} \quad \text{在 } \Omega \text{ 内},$$
$$\boldsymbol{u} = \boldsymbol{0} \quad \text{在 } \Gamma_0 \text{ 上},$$
$$\boldsymbol{\sigma}(\boldsymbol{u})\boldsymbol{\nu} = \boldsymbol{g} \quad \text{在 } \Gamma_1 \text{ 上},$$

其中

$$\boldsymbol{\sigma}(\boldsymbol{u}) := \lambda(\mathrm{tr}\,\boldsymbol{e}(\boldsymbol{u}))\boldsymbol{I} + 2\mu\boldsymbol{e}(\boldsymbol{u}),$$

称为三维线性化弹性的边值问题. 如同 $\mathbb{R}^3$ 中的 Stokes 方程 (6.14 节), 它又提供了一个含有三个未知函数、三个偏微分方程的方程组的实例.

更明确地说, 上述边值问题是下述物理现象的数学模型: 集合 $\overline{\Omega}$ 是弹性体的参考构形[78], 它承受作用在其内部的外加体积力, 其单位体积密度为 $\boldsymbol{f} = (f_i) : \Omega \to \mathbb{R}^3$; 又承受作用在集合 Ω 边界 Γ 的一部分 Γ_1 上的外加面力, 其单位面积密度为 $\boldsymbol{g} = (g_i) : \Gamma_1 \to \mathbb{R}^3$.

上述问题的未知量是位移向量场 $\boldsymbol{u} = (u_i) : \overline{\Omega} \to \mathbb{R}^3$, 即向量 $\boldsymbol{u}(x) = (u_i(x))$ 表示参考构形 $\overline{\Omega}$ 的每一点 x 在外加力的作用下所产生的位移 (图 9.7-1). 弹性体被假定在边界的另一部分 $\Gamma_0 := \Gamma - \Gamma_1$ 上满足关于位置的齐次边界条件. 这意味着未知位移向量场在 Γ_0 上满足边界条件 $\boldsymbol{u} = \boldsymbol{0}$.

最后, 假定构成物体的弹性材料是齐次的、各向同性的, 而且参考构形 $\overline{\Omega}$ 处于自然状态. 这些假定意味着, "在一阶的范围内", 材料的变化行为只由两个常数 λ 及 μ 控制, 它们称为材料的 Lamé 常数[79]. 实验证据说明, 现实弹性材料的 Lamé 常数满足不等式 $\lambda \geqslant 0$ 及 $\mu > 0$, 因此在定理 6.16-1 中, 已经假定其成立 (Lamé 常数用以衡量构成材料的 "刚度": 它们越大, 材料的刚度越大).

对称矩阵场 $\boldsymbol{e}(\boldsymbol{u}) = (e_{ij}(\boldsymbol{u})) : \overline{\Omega} \to \mathbb{S}^3$ 称为线性化应变张量场, 对称矩阵场 $\boldsymbol{\sigma}(\boldsymbol{u}) = (\sigma_{ij}(\boldsymbol{u})) : \overline{\Omega} \to \mathbb{S}^3$ 称为线性化应力张量场, 而它们的分量 $e_{ij}(\boldsymbol{u})$ 及 $\sigma_{ij}(\boldsymbol{u})$ 则分别称线性化应变及线性化应力. 这两个线性化张量之间的线性关系 $\boldsymbol{\sigma}(\boldsymbol{u}) = \lambda(\mathrm{tr}\,\boldsymbol{e}(\boldsymbol{u})\boldsymbol{I} + 2\mu\boldsymbol{e}(\boldsymbol{u}))$ 在弹性理论中被称为 Hooke 定律[80], 它刻画出齐次且各向同性线性弹性体的本构特征.

出现在定理 6.16-1 中的泛函 $J : \boldsymbol{v} \in \boldsymbol{V} \to \mathbb{R}$ 表示齐次、各向同性的线性弹性体的能量, 而定理 6.16-1 中的变分方程 $a(\boldsymbol{u}, \boldsymbol{v}) = \ell(\boldsymbol{v})$ 对所有 $\boldsymbol{v} \in \boldsymbol{V}$ 则构成线性化虚功原理, 因此它对所有运动学上容许的位移 $\boldsymbol{v} \in \boldsymbol{V}$, 也就是所有在 Γ_0 上满足边界条件 $\boldsymbol{v} = \boldsymbol{0}$ 的向量场 $\boldsymbol{v} \in \boldsymbol{V}$ 成立.

---

78) 在此用到的所有关于弹性理论概念 (如参考构形、弹性体、作用力、静荷载等) 的详尽介绍可参阅如 CIARLET [1988].

79) 冠名源自 Gabriel Lamé (1795—1870).

80) 冠名源自 Robert Hooke (1635—1703).

注 非齐次各向异性的线性弹性体的能量采取更一般的形式

$$J(\boldsymbol{v})=\frac{1}{2}\int_{\Omega}\boldsymbol{A}\boldsymbol{e}(\boldsymbol{v}):\boldsymbol{e}(\boldsymbol{v})dx-\left(\int_{\Omega}\boldsymbol{f}\cdot\boldsymbol{v}dx+\int_{\Gamma_1}\boldsymbol{g}\cdot\boldsymbol{v}d\Gamma\right),\quad \text{对所有 } \boldsymbol{v}\in\boldsymbol{V},$$

其中弹性张量 $\boldsymbol{A}=(A_{ijk\ell})$ 具有下述性质: 其分量 $A_{ijk\ell}$ 属于 $L^{\infty}(\Omega)$, 它们满足对称性 $A_{ijk\ell}=A_{jik\ell}=A_{k\ell ij}$, 且存在常数 $\alpha>0$ 使得

$$\boldsymbol{A}(x)\boldsymbol{t}:\boldsymbol{t}\geqslant\alpha\boldsymbol{t}:\boldsymbol{t}\quad \text{对几乎所有 } x\in\Omega \text{ 及所有矩阵 } \boldsymbol{t}=(t_{ij})\in\mathbb{S}^3,$$

其中 $(\boldsymbol{A}(x)\boldsymbol{t})_{ij}:=A_{ijk\ell}(x)t_{k\ell}$. 在这种情况下, 关系式 $\boldsymbol{\sigma}(\boldsymbol{u})=\lambda(\mathrm{tr}\,\boldsymbol{e}(\boldsymbol{u}))\boldsymbol{I}+2\mu\boldsymbol{e}(\boldsymbol{u})$ 由下面的更一般的线性关系取而代之:

$$\boldsymbol{\sigma}(\boldsymbol{u})=\boldsymbol{A}\boldsymbol{e}(\boldsymbol{u}).$$

对齐次各向同性的线性弹性体的特殊情况 (这相应于 Hooke 定律), 相应地有

$$A_{ijk\ell}=\lambda\delta_{ij}\delta_{k\ell}+\mu(\delta_{ik}\delta_{j\ell}+\delta_{i\ell}\delta_{jk}).$$

如果 $d\Gamma\text{-meas-}\Gamma_0>0$ 且 $d\Gamma\text{-meas-}\Gamma_1>0$, 线性化弹性的边值问题就称为*位移–牵引问题*; 若 $\Gamma_0=\Gamma$, 则称其为*纯位移问题*; 若 $\Gamma_1=\Gamma$, 则称其为*纯牵引问题* (对后一问题的讨论并未包含在定理 6.16-1 中, 而是习题 6.16-2 中的主题).

按照 6.13 节中的框架, 可以给出纯位移问题的混合及对偶形式, 见习题 6.16-3.

可以证明, 如果 $\Gamma=\Gamma_0$ 而且问题的资料 (Ω 的边界及右端项 $\boldsymbol{f}$) 具有附加的正则性, 则定理 6.16-1 中给出在此情况下属于空间 $\boldsymbol{V}=\boldsymbol{H}_0^1(\Omega)$ 的弱解也具有附加的正则性.

定理 6.16-2 (线性化弹性纯位移问题弱解的正则性[81]) *设 Ω 是 $\mathbb{R}^3$ 中具有 $\mathcal{C}^2$ 类边界 Γ 的区域. 又设出现在定理 6.16-1 中的 $\boldsymbol{f}\in\boldsymbol{L}^p(\Omega), p\geqslant 6/5$, 而 $\Gamma_0:=\Gamma$. 则在这种情况下, 弱解 $\boldsymbol{u}\in\boldsymbol{H}_0^1(\Omega)$ 属于空间 $\boldsymbol{W}^{2,p}(\Omega)$ 且满足*

$$-\mathbf{div}\{\lambda(\mathrm{tr}\,\boldsymbol{e}(\boldsymbol{u}))\boldsymbol{I}+2\mu\boldsymbol{e}(\boldsymbol{u})\}=\boldsymbol{f}\quad \text{在 } \boldsymbol{L}^p(\Omega) \text{ 中}.$$

习题

6.16-1 下面 $N\geqslant 2$ 是整数.

(1) 设 λ 和 μ 是两个常数, 满足 $N\lambda+2\mu>0$ 及 $\mu>0$. 证明存在常数 $\alpha=\alpha(N,\lambda,\mu)>0$ 使得

$$\lambda(\mathrm{tr}\,\boldsymbol{B})^2+2\mu\mathrm{tr}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B})\geqslant\alpha\mathrm{tr}(\boldsymbol{B}^{\mathrm{T}}\boldsymbol{B}),\quad \text{对所有 } \boldsymbol{B}\in\mathbb{M}^N.$$

因此, 这个不等式在 $N=3$ 的特殊情况意味着如果 $3\lambda+2\mu>0$ 且 $\mu>0$, 定理 6.16-1 中的双线性形式仍保持 $\boldsymbol{V}$ 强制性.

[81] 该定理的证明很长、需一些精细的处理手法, 其梗概可见 Ciarlet [1988, 6.3 节].

(2) 反之, 设 λ 和 μ 是两个常数, 具有以下性质: (1) 中的不等式对某个常数 $\alpha > 0$ 满足. 证明, 必然成立 $N\lambda + 2\mu > 0$ 及 $\mu > 0$.

6.16-2 这个习题将定理 6.16-1 中的存在性和唯一性结果推广到三维线性弹性的纯牵引问题.

设 Ω 是 $\mathbb{R}^3$ 中具有边界 Γ 的区域, 又设常数 $\lambda \geqslant 0, \mu > 0$ 及向量场 $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega), \boldsymbol{g} \in \boldsymbol{L}^2(\Gamma)$ 给定. 证明下列极小化问题: 求 $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ 使得 $J(\boldsymbol{u}) = \inf_{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)} J(\boldsymbol{v})$, 其中

$$
\begin{aligned}
J(\boldsymbol{v}) := & \frac{1}{2} \int_\Omega \{\lambda (\operatorname{tr} \boldsymbol{e}(\boldsymbol{v}))^2 + 2\mu \boldsymbol{e}(\boldsymbol{v}) : \boldsymbol{e}(\boldsymbol{v})\} dx \\
& - \left(\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} dx + \int_\Gamma \boldsymbol{g} \cdot \boldsymbol{v} d\Gamma \right), \quad \text{对所有 } \boldsymbol{v} \in \boldsymbol{H}^1(\Omega),
\end{aligned}
$$

有解的充分必要条件是

$$
\int_\Omega \boldsymbol{f} \cdot \boldsymbol{v} dx + \int_\Gamma \boldsymbol{g} \cdot \boldsymbol{v} d\Gamma = 0, \quad \text{对所有 } \boldsymbol{v} \in \mathbf{Ker}\, \boldsymbol{\nabla}_s,
$$

即对所有无穷小刚体位移 (6.15 节), 而且这个解在相差一个无穷小刚体位移的情况下是唯一的.

提示: 用定理 6.15-3.

6.16-3 本题中的问题 (3) 和 (5), 对于线性化弹性的纯位移问题, 即定理 6.16-1 中 $\Gamma_0 = \Gamma$ 这一特殊情况, 分别给出其一个混合形式[82] 和一个对偶形式. 假设与符号也与该定理中的相同.

(1) 定义空间

$$
\mathbb{H}(\mathbf{div}; \Omega) := \{\boldsymbol{\tau} \in \mathbb{L}^2(\Omega); \mathbf{div}\, \boldsymbol{\tau} \in \boldsymbol{L}^2(\Omega)\}.
$$

证明, 装备如下定义的范数:

$$
\|\boldsymbol{\tau}\|_{\mathbb{H}(\mathbf{div};\Omega)} := (\|\boldsymbol{\tau}\|_{0,\Omega}^2 + \|\mathbf{div}\, \boldsymbol{\tau}\|_{0,\Omega}^2)^{1/2}, \quad \text{对每个 } \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}; \Omega),
$$

空间 $\mathbb{H}(\mathbf{div}; \Omega)$ 是 Hilbert 空间.

(2) 证明由下式定义的映射 $\boldsymbol{B} : \mathbb{S}^3 \to \mathbb{S}^3$,

$$
\boldsymbol{B\sigma} := \frac{1}{2\mu} \left(\boldsymbol{\sigma} - \frac{\lambda}{3\lambda + 2\mu} (\operatorname{tr} \boldsymbol{\sigma}) \boldsymbol{I} \right), \quad \text{对每个 } \boldsymbol{\sigma} \in \mathbb{S}^3,
$$

是由

$$
\boldsymbol{Ae} := \lambda (\operatorname{tr} \boldsymbol{e}) \boldsymbol{I} + 2\mu \boldsymbol{e}, \quad \text{对每个 } \boldsymbol{e} \in \mathbb{S}^3
$$

定义的映射 $\boldsymbol{A} : \mathbb{S}^3 \to \mathbb{S}^3$ 之逆.

(3) 证明存在唯一偶对 $(\boldsymbol{\sigma}, \boldsymbol{\lambda}) \in \mathbb{H}(\mathbf{div}; \Omega) \times \boldsymbol{L}^2(\Omega)$ 满足:

$$
\begin{aligned}
& \int_\Omega \boldsymbol{B\sigma} : \boldsymbol{\tau} dx + \int_\Omega \mathbf{div} \boldsymbol{\tau} \cdot \boldsymbol{\lambda} dx = 0, \quad \text{对所有 } \boldsymbol{\tau} \in \mathbb{H}(\mathbf{div}; \Omega), \\
& \int_\Omega \mathbf{div} \boldsymbol{\sigma} \cdot \boldsymbol{\mu} dx = -\int_\Omega \boldsymbol{f} \cdot \boldsymbol{\mu} dx, \quad \text{对所有 } \boldsymbol{\mu} \in \boldsymbol{L}^2(\Omega).
\end{aligned}
$$

[82] 另一种混合形式也是可能的; 例如, 可见下文第 11 节:

D. N. ARNOLD; R. S. FALK; R. WINTHER [2006]: Finite element exterior calculus, homological techniques, and applications. Acta Numerica, Volume 15 (A. ISERLES, editor), pp. 1–155, Cambridge University Press, Cambridge, UK.

(4) 设 $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ 是下述二次极小问题的唯一解 (定理 6.16-1): 求 $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ 使得 $J(\boldsymbol{u}) = \inf_{\boldsymbol{v}\in \boldsymbol{H}_0^1(\Omega)} J(\boldsymbol{v})$, 其中

$$J(\boldsymbol{v}) := \frac{1}{2}\int_\Omega \{\lambda(\operatorname{tr} \boldsymbol{e}(\boldsymbol{v}))^2 + 2\mu \boldsymbol{e}(\boldsymbol{v}) : \boldsymbol{e}(\boldsymbol{v})\}dx - \int_\Omega \boldsymbol{f}\cdot \boldsymbol{v}dx, \quad \text{对每个 } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega).$$

证明

$$\boldsymbol{\sigma} = \boldsymbol{A}\boldsymbol{e}(\boldsymbol{u}) \quad \text{及} \quad \boldsymbol{\lambda} = \boldsymbol{u}.$$

(5) 证明矩阵场 $\boldsymbol{\sigma} = \boldsymbol{A}\boldsymbol{e}(\boldsymbol{u})$ 是下述约束二次极小化问题的唯一解:

$$\boldsymbol{\sigma} \in \mathbb{U}_{\boldsymbol{f}} := \{\boldsymbol{\tau} \in \mathbb{H}(\mathbf{div};\Omega); \mathbf{div}\,\boldsymbol{\tau} + \boldsymbol{f} = \boldsymbol{0} \text{ 在 } \boldsymbol{L}^2(\Omega) \text{ 中}\},$$
$$I(\boldsymbol{\sigma}) = \inf_{\boldsymbol{\tau}\in\mathbb{U}_{\boldsymbol{f}}} \boldsymbol{I}(\boldsymbol{\tau}), \quad \text{其中 } \boldsymbol{I}(\boldsymbol{\tau}) := \frac{1}{2}\int_\Omega B\boldsymbol{\tau} : \boldsymbol{\tau}dx, \quad \text{对每个 } \boldsymbol{\tau} \in \mathbb{L}_s^2(\Omega).$$

提示: 仿照定理 6.13-2 的证明.

注　在弹性理论中, $\mathbb{U}_{\boldsymbol{f}}$ 称为容许应力集合, 而泛函 $I : \mathbb{L}^2(\Omega) \to \mathbb{R}$ 则称为余能.

6.16-4　希腊和拉丁指标分别在集合 $\{1,2\}$ 和 $\{1,2,3\}$ 中变化, 对于重复指标, 仍用其求和约定. 设 Ω 为 $\mathbb{R}^2$ 中的区域, Γ_1 是 $\Gamma := \partial\Omega$ 的相对开子集使得 $d\Gamma\text{-meas-}\Gamma_0 > 0$, 其中 $\Gamma_0 := \Gamma - \Gamma_1$. 又设 $\boldsymbol{f} = (f_i) \in L^2(\Omega)$ 是给定的向量场. 定义空间

$$\boldsymbol{V} := \{\boldsymbol{v} = (v_i) \in H^1(\Omega) \times H^1(\Omega) \times H^2(\Omega); v_i = \partial_\nu v_3 = 0 \text{ 在 } \Gamma_0 \text{ 上}\},$$

及泛函 $\boldsymbol{J} : \boldsymbol{V} \to \mathbb{R}$ 如下:

$$J(\boldsymbol{v}) := \frac{1}{2}\int_\Omega \left\{\frac{\varepsilon^3}{3} a_{\alpha\beta\sigma\tau}\partial_{\sigma\tau}v_3\partial_{\alpha\beta}v_3 + \varepsilon a_{\alpha\beta\sigma\tau}e_{\sigma\tau}(\boldsymbol{v})e_{\alpha\beta}(\boldsymbol{v})\right\}dx - \int_\Omega f_i v_i dx, \quad \boldsymbol{v} = (v_i) \in \boldsymbol{V},$$

其中 $\varepsilon > 0$ 是常数, $a_{\alpha\beta\sigma\tau} = \alpha_{\beta\alpha\sigma\tau} = a_{\sigma\tau\alpha\beta}$ 是具有下述性质的常数: 存在常数 $C > 0$ 使得

$$a_{\alpha\beta\sigma\tau}t_{\sigma\tau}t_{\alpha\beta} \geqslant Ct_{\alpha\beta}t_{\alpha\beta} \quad \text{对所有 } (t_{\alpha\beta}) \in \mathbb{S}^2,$$

而

$$e_{\alpha\beta}(\boldsymbol{v}) := \frac{1}{2}(\partial_\alpha v_\beta + \partial_\beta v_\alpha).$$

(1) 证明存在唯一的向量场 $\boldsymbol{u} \in \boldsymbol{V}$ 使得 $J(\boldsymbol{u}) = \inf_{\boldsymbol{v}\in\boldsymbol{V}} J(\boldsymbol{v})$.

(2) 假设 $\boldsymbol{u} = (u_i) \in H^2(\Omega) \times H^2(\Omega) \times H^4(\Omega)$. 证明 $\boldsymbol{u}$ 满足下述边值问题:

$$\partial_{\alpha\beta}m_{\alpha\beta}(\boldsymbol{u}) = f_3 \quad \text{和} \quad -\partial_\beta n_{\alpha\beta}(\boldsymbol{u}) = f_\alpha \quad \text{在 } \Omega \text{ 中},$$
$$u_i = \partial_\nu u_3 = 0 \quad \text{在 } \Gamma_0 \text{ 上},$$
$$m_{\alpha\beta}(\boldsymbol{u})\nu_\alpha\nu_\beta = 0, n_{\alpha\beta}(\boldsymbol{u})\nu_\beta = 0 \quad \text{和} \quad (\partial_\alpha m_{\alpha\beta}(\boldsymbol{u}))\nu_\beta + \partial_\tau(m_{\alpha\beta}(\boldsymbol{u})\nu_\alpha\tau_\beta) = 0 \text{ 在 } \Gamma_1 \text{ 上},$$

其中

$$m_{\alpha\beta}(\boldsymbol{u}) := \frac{\varepsilon^3}{3}a_{\alpha\beta\sigma\tau}\partial_{\sigma\tau}u_3 \quad \text{和} \quad n_{\alpha\beta}(\boldsymbol{u}) := \varepsilon a_{\alpha\beta\sigma\tau}e_{\sigma\tau}(\boldsymbol{u}).$$

上述边值问题构成关于厚度为 2ε, 夹紧在边界的一部分 Γ_0 上的线性弹性板的 Kirchhoff-Love 理论[83] 方程组. 要注意, 这个问题实际上由两组独立的边值问题组成, 一组 (以 u_3 为未知函数) 构成弯曲方程 (在定理 6.8-7 中已碰到过, 只是用的符号不同), 而另一组 (以 u_1, u_2 为未知函数) 构成薄膜方程.

注 常数 $a_{\alpha\beta\sigma\tau}$ 是板的弹性张量的分量. 它们可以用构成板的弹性材料的 Lamé 常数 $\lambda \geqslant 0$ 和 $\mu > 0$ 来表示:

$$a_{\alpha\beta\sigma\tau} = \frac{4\lambda\mu}{(\lambda+2\mu)}\delta_{\alpha\beta}\delta_{\sigma\tau} + 2\mu(\delta_{\alpha\sigma}\delta_{\beta\sigma} + \delta_{\alpha\tau}\delta_{\beta\sigma}).$$

6.17 经典 Poincaré 引理, 及其作为 J. L. Lions 引理和 △ 次椭圆性应用的弱形式

整个这一节中, 对于重复指标仍用求和约定. 给定 $\mathbb{R}^N$ 的一个开子集 Ω, 考察由下式定义的线性算子 $\mathbf{grad} : \mathcal{C}^2(\Omega) \to \mathcal{C}^1(\Omega; \mathbb{R}^N)$:

$$p \in \mathcal{C}^2(\Omega) \to \mathbf{grad}\, p := (\partial_i p) \in \boldsymbol{\mathcal{C}}^1(\Omega) := \mathcal{C}^1(\Omega; \mathbb{R}^N).$$

由此产生一个自然的问题, 这个线性算子是否可逆, 即给定一个向量场 $\boldsymbol{h} = (h_i) \in \mathcal{C}^1(\Omega; \mathbb{R}^N)$, 是否存在一个函数 $p \in \mathcal{C}^2(\Omega)$ 使得

$$\mathbf{grad}\, p = \boldsymbol{h} \quad 在\ \Omega\ 中,$$

或等价地使用

$$\partial_i p = h_i \quad 在\ \Omega\ 中, \quad 1 \leqslant i \leqslant N.$$

如果这个情况成立, 则应有 $\partial_{ij} p = \partial_{ji} p$, 故显然函数 h_i 必须满足相容性条件

$$\partial_i h_j - \partial_j h_i = 0 \quad 在\ \mathcal{C}(\Omega)\ 中, \quad 1 \leqslant i, j \leqslant N,$$

或等价地

$$\mathbf{curl}\, \boldsymbol{h} = \mathbf{0} \quad 在\ \boldsymbol{\mathcal{C}}(\Omega) := \mathcal{C}(\Omega; \mathbb{R}^N)\ 中,$$

其中对任意整数 $N \geqslant 2$ 旋度算子 $\mathbf{curl} : \mathcal{C}^1(\Omega; \mathbb{R}^N) \to \mathcal{C}(\Omega; \mathbb{R}^{\frac{N(N-1)}{2}})$ 定义为[84]

$$(\mathbf{curl}\, \boldsymbol{h})_{ij} := (\partial_i h_j - \partial_j h_i), \quad 1 \leqslant i, j \leqslant N, \quad 对每个\ \boldsymbol{h} \in \mathcal{C}^1(\Omega; \mathbb{R}^N).$$

值得注意的是, 如果开集 Ω 是单连通的, 这些必要条件就变成充分的了. 这就是下面定理 6.17-2 中确立的经典 Poincaré 引理的实质 (“经典”, 是相对于定理 6.17-4 中确立的这个引理的 “弱” 形式而言的).

---

[83] 这些方程在 CIARLET [1997, 第 1 章] 中有详尽的讨论.

[84] 这个定义的合理性可参阅如:

G. CSATO; B. DACOROGNA; O. KNEUSS [2011]: The Pullback Equation. Birkhäuser, Basel.

在证明该引理之前, 我们首先建立一个技术上的, 而其本身也是很有意义的结果: 尽管路径, 或者同伦, 当它们被用来定义一般单连通拓扑空间 X 时, 只假定分别是从 $[0,1]$ 或 $[0,1]\times[0,1]$ 到 X 中的连续映射 (1.9 节), 但当 X 是 $\mathbb{R}^N$ 的开子集时, 可以假定它们是 $\mathcal{C}^\infty$ 类的 (实际上只要这些映射是 $\mathcal{C}^2$ 类的, 对我们的目的而言已经够了, 但证明它们是 $\mathcal{C}^\infty$ 类的, 并未增加额外工作).

定理 6.17-1 设 Ω 是 $\mathbb{R}^N$ 的单连通开子集. 则:

(a) 给定任意两个不同的点 $x\in\Omega$ 和 $y\in\Omega$, 在 Ω 中存在一个路径 $\gamma\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 联结 x 与 y.

(b) 给定任意两个不同点 $x\in\Omega$ 和 $y\in\Omega$, 设 $\gamma^0\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 和 $\gamma^1\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 是在 Ω 中任意两条联结 x 与 y 的不同路径. 则在 Ω 中存在一个同伦 $\boldsymbol{H}\in\mathcal{C}^\infty([0,1]\times[0,1];\mathbb{R}^N)$ 联结 γ^0 与 γ^1.

证明 (i) 因为单连通空间是弧连通的, 在 Ω 中存在一个路径 $\boldsymbol{\pi}\in\mathcal{C}([0,1];\mathbb{R})$ 联结 x 与 y. 因为 $\mathrm{Im}\,\boldsymbol{\pi}$ 是 Ω 的紧子集, 故存在 $\delta>0$ 使得

$$\bigcup_{x\in\mathrm{Im}\,\boldsymbol{\pi}}\overline{B(x;\delta)}\subset\Omega.$$

设 $\tilde{\boldsymbol{\pi}}=(\tilde{\pi}_i)_{i=1}^N\in\mathcal{C}(\mathbb{R};\mathbb{R}^N)$ 是 $\boldsymbol{\pi}$ 的一个延拓 (根据 Tietze-Urysohn 定理这样的延拓是存在的, 见定理 1.7-7), 即使得 $\tilde{\boldsymbol{\pi}}|_{[0,1]}=\boldsymbol{\pi}$. 又设 $(\tilde{\pi}_{i,\varepsilon})_{\varepsilon>0}, 1\leqslant i\leqslant N$, 是每一个分量 $\tilde{\pi}_i$ 的正则化族 (2.6 节), 令 $\tilde{\boldsymbol{\pi}}_\varepsilon:=(\tilde{\boldsymbol{\pi}}_{i,\varepsilon})_{i=1}^N, \varepsilon>0$. 则 $\tilde{\boldsymbol{\pi}}_\varepsilon\in\mathcal{C}^\infty(\mathbb{R};\mathbb{R}^N)$ 且

$$\sup_{0\leqslant t\leqslant 1}|\tilde{\boldsymbol{\pi}}_\varepsilon(\ell)-\boldsymbol{\pi}(t)|\leqslant\frac{\delta}{2}\quad\text{对充分小的 }\varepsilon>0$$

(定理 2.6-1(b)). 固定这个 $\varepsilon>0$ 并设 $\boldsymbol{\gamma}:[0,1]\to\mathbb{R}^N$ 由

$$\boldsymbol{\gamma}(t):=\tilde{\boldsymbol{\pi}}_\varepsilon(t)+(1-t)(x-\tilde{\boldsymbol{\pi}}_\varepsilon(0))+t(y-\tilde{\boldsymbol{\pi}}_\varepsilon(1)),\quad 0\leqslant t\leqslant 1,$$

定义. 则 $\boldsymbol{\gamma}\in\mathcal{C}^\infty([0,1];\mathbb{R}^N), \gamma(0)=x, \gamma(1)=y$, 且

$$\begin{aligned}|\boldsymbol{\gamma}(t)-\boldsymbol{\pi}(t)|&\leqslant|\tilde{\boldsymbol{\pi}}_\varepsilon(t)-\boldsymbol{\pi}(t)|+(1-t)|\boldsymbol{\pi}(0)-\tilde{\boldsymbol{\pi}}_\varepsilon(0)|+t|\boldsymbol{\pi}(1)-\tilde{\boldsymbol{\pi}}_\varepsilon(1)|\\&\leqslant\frac{\delta}{2}+(1-t)\frac{\delta}{2}+t\frac{\delta}{2}=\delta,\quad 0\leqslant t\leqslant 1,\end{aligned}$$

此因 $\boldsymbol{\pi}(0)=x, \boldsymbol{\pi}(1)=y$, 这样就有 $\boldsymbol{\gamma}(t)\in\cup_{x\in\mathrm{Im}\,\boldsymbol{\pi}}\overline{B(x;\delta)}\subset\Omega$ 对所有 $t\in[0,1]$.

(ii) 根据单连通集的定义, 在 Ω 中存在一个同伦 $\boldsymbol{G}\in\mathcal{C}([0,1]\times[0,1];\mathbb{R}^N)$ 联结 γ^0 与 γ^1. 因为 $\mathrm{Im}\,\boldsymbol{G}$ 是 Ω 的紧子集, 故存在 $\delta>0$ 使得

$$\bigcup_{x\in\mathrm{Im}\,\boldsymbol{G}}\overline{B(x;\delta)}\subset\Omega.$$

设 $\tilde{\boldsymbol{G}}=(\tilde{G}_i)_{i=1}^N\in\mathcal{C}(\mathbb{R}\times\mathbb{R};\mathbb{R}^N)$ 是 $\boldsymbol{G}$ 的一个延拓 (还是由 Tietze-Urysohn 定理知其存在), 即使得 $\tilde{\boldsymbol{G}}|_{[0,1]\times[0,1]}=\boldsymbol{G}$, 又设 $(\tilde{G}_{i,\varepsilon})_{\varepsilon>0}$ 是每一个分量 $\tilde{G}_i, 1\leqslant i\leqslant N$,

的正则化族, 令 $\tilde{\boldsymbol{G}}_\varepsilon := (\tilde{G}_{i,\varepsilon})_{i=1}^N, \varepsilon > 0$. 则 $\tilde{\boldsymbol{G}}_\varepsilon \in \mathcal{C}^\infty(\mathbb{R}\times\mathbb{R};\mathbb{R}^N)$ 且 (仍根据定理 2.6-1(b))

$$\sup_{0\leqslant t\leqslant 1, 0\leqslant\lambda\leqslant 1} |\tilde{\boldsymbol{G}}_\varepsilon(t,\lambda) - \boldsymbol{G}(t,\lambda)| \leqslant \frac{\delta}{2} \quad \text{对充分小的 } \varepsilon > 0.$$

固定这个 $\varepsilon > 0$ 并设 $\boldsymbol{H} : [0,1]\times[0,1] \to \mathbb{R}^N$ 由

$$\boldsymbol{H}(t,\lambda) := \tilde{\boldsymbol{G}}_\varepsilon(t,\lambda) + (1-\lambda)(\boldsymbol{\gamma}^0(t) - \tilde{\boldsymbol{G}}_\varepsilon(t,0)) + \lambda(\boldsymbol{\gamma}^1(t) - \tilde{\boldsymbol{G}}_\varepsilon(t,1))$$

定义. 则 $\boldsymbol{H} \in \mathcal{C}^\infty([0,1]\times[0,1];\mathbb{R}^N), \boldsymbol{H}(t,0) = \boldsymbol{\gamma}^0(t) = \boldsymbol{G}(t,0)$ 及 $\boldsymbol{H}(t,1) = \boldsymbol{\gamma}^1(t) = \boldsymbol{G}(t,1)$ 对 $0\leqslant t\leqslant 1$, 且

$$\begin{aligned}|\boldsymbol{H}(t,\lambda) - \boldsymbol{G}(t,\lambda)| \leqslant{}& |\tilde{\boldsymbol{G}}_\varepsilon(t,\lambda) - \boldsymbol{G}(t,\lambda)| \\ &+(1-\lambda)|\boldsymbol{G}(t,0) - \tilde{\boldsymbol{G}}_\varepsilon(t,0)| + \lambda|\boldsymbol{G}(t,1) - \tilde{\boldsymbol{G}}_\varepsilon(t,1)| \\ \leqslant{}& \frac{\delta}{2} + (1-\lambda)\frac{\delta}{2} + \lambda\frac{\delta}{2} = \delta, \quad 0\leqslant t\leqslant 1, 0\leqslant\lambda\leqslant 1.\end{aligned}$$

所以 $\boldsymbol{H}(t,\lambda) \in \cup_{x\in\operatorname{Im}\boldsymbol{G}}\overline{B(x;\delta)} \subset \Omega$ 对所有 $0\leqslant t\leqslant 1, 0\leqslant\lambda\leqslant 1$. □

我们现在证明 "经典" Poincaré 引理. 在下面的证明中, 拉丁指标在集合 $\{1,2,\ldots,N\}$ 中变化, 而关于重复指标的求和约定, 也按此进行.

定理 6.17-2 (经典 Poincaré 引理; 或 $\mathcal{C}^2(\Omega)$ 中的 Poincaré 引理[85]) *设 Ω 是 $\mathbb{R}^N$ 中的单连通开子集, 又设给定一个向量场 $\boldsymbol{h} \in \mathcal{C}^1(\Omega)$, 它满足*

$$\mathbf{curl}\,\boldsymbol{h} = \mathbf{0} \quad \text{在 } \Omega \text{ 中}.$$

则存在函数 $p \in \mathcal{C}^2(\Omega)$ 使得

$$\mathbf{grad}\,p = \boldsymbol{h} \quad \text{在 } \Omega \text{ 中}.$$

而且方程 $\mathbf{grad}\,\tilde{p} = \boldsymbol{h}$ 在 Ω 中的任何一个解均可表示为以下形式: $\tilde{p} = p + C$ 对某个常数 C.

证明 (i) 设 $x^0 \in \Omega$ 是给定的点. 因为 Ω 也是弧连通的, 给定不同于 x^0 的任一点 $x^1 \in \Omega$, 在 Ω 中存在路径 $\gamma = (\gamma_i) \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 联结 x^0 及 x^1 (定理 6.17-1), 即 $\gamma(0) = x^0, \gamma(1) = x^1$, 且 $\gamma(t) \in \Omega$ 对所有 $0\leqslant t\leqslant 1$.

[85] 冠名源自 Henri Poincaré, 他在 1886 年实际上已经提及这个结果的一个 (关于任意阶微分形式的) 推广, 其证明是随后 1889 年由 Vito Volterra 给出的. 但是在此陈述的 "Poincaré 引理" (即对于一阶微分形式的) 实际上 (对 $N = 2$) 可追溯到 Alexis Claude de Clairaut, Leonhard Euler 及 Alexi Fontaine des Bertins. 他们在 1740 年前后独立地证明了这一结果. 关于这个引理的产生及其推广的更多细节, 读者可参阅:

H. Samelson [2001]: Differential forms, the early days; or the stories of Deahna's theorem and of Volterra's theorem. American Mathematical Monthly 108, 552–530.

关于 Henri Poincaré 杰出贡献的高度概括可见 Gray [2012].

设给定一向量场 $\boldsymbol{h} = (h_i) \in \boldsymbol{\mathcal{C}}^1(\Omega)$. 如果存在函数 $p \in \mathcal{C}^2(\Omega)$ 并满足 $\partial_i p = h_i$ 在 Ω 中, $1 \leqslant i \leqslant N$, 则由 $P(t) := p(\gamma(t)), 0 \leqslant t \leqslant 1$, 定义的函数 $P \in \mathcal{C}^1[0,1]$ 依赖于事先指定的路径 γ, 并满足

$$\frac{dP}{dt}(t) = \partial_i p(\gamma(t))\frac{d\gamma_i}{dt}(t) = h_i(\gamma(t))\frac{d\gamma_i}{dt}(t), \quad 0 \leqslant t \leqslant 1.$$

据此考虑, 我们首先注意到, 对任意的 $P^0 \in \mathbb{R}$, 存在唯一仍依赖路径 γ 的线性 Cauchy 问题:

$$\frac{dP}{dt}(t) = h_i(\gamma(t))\frac{d\gamma_i}{dt}(t), \quad 0 \leqslant t \leqslant 1, \text{ 且 } P(0) = P^0$$

的解 (定理 3.8-2). 这个结果顺便还证明了, 如果方程组

$$\partial_i p(x) = h_i(x) \quad \text{在 } \Omega \text{ 中}, \quad 1 \leqslant i \leqslant N, \text{ 且 } p(x^0) = P^0$$

有解, 则这个解是唯一的.

(ii) 要使得 (i) 中 Cauchy 问题解的值 $P(1)$ 可以成为未知函数值 $p(x^1)$, 当然数值 $P(1)$ 必须与所选联结 x^0 和 x^1 的路径无关. 我们将证明, 这一性质取决于函数 h_i 所满足的相容性条件 $\partial_j h_i = \partial_i h_j$, 以及 Ω 是单连通的这一假设.

设 $\gamma_0 \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 和 $\gamma_1 \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 是 Ω 中联结 x^0 与 x^1 的两个路径. 因 Ω 是单连通的, 故在 Ω 中存在联结 γ^0 与 γ^1 的同伦 $\boldsymbol{G} = (G_i)_{i=1}^N \in \mathcal{C}^\infty([0,1] \times [0,1];\mathbb{R}^N)$ (定理 6.17-1), 即使得

$$\begin{gathered}
\boldsymbol{G}(t,\lambda) \in \Omega \quad \text{对所有 } 0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1, \\
\boldsymbol{G}(\cdot,0) = \gamma_0 \quad \text{及} \quad \boldsymbol{G}(\cdot,1) = \gamma_1, \\
\boldsymbol{G}(0,\lambda) = x^0 \quad \text{及} \quad \boldsymbol{G}(1,\lambda) = x^1 \quad \text{对所有 } 0 \leqslant \lambda \leqslant 1.
\end{gathered}$$

对每个 $\lambda \in [0,1]$, 设 $P(\cdot,\lambda) \in \mathcal{C}^1[0,1]$ 表示相应于在 Ω 中联结 x^0 与 x^1 的特定路径 $\boldsymbol{G}(\cdot,\lambda) = (\boldsymbol{G}_i(\cdot,\lambda))$ 的 Cauchy 问题的唯一解. 我们就有

$$\frac{\partial P}{\partial t}(t,\lambda) = h_i(\boldsymbol{G}(t,\lambda))\frac{\partial G_i}{\partial t}(t,\lambda), \quad 0 \leqslant t \leqslant 1 \quad \text{及} \quad P(0,\lambda) = P^0, \quad 0 \leqslant \lambda \leqslant 1.$$

我们的目标在于证明

$$\frac{\partial P}{\partial \lambda}(1,\lambda) = 0, \quad 0 \leqslant \lambda \leqslant 1$$

(易见 $P \in \mathcal{C}^1([0,1] \times [0,1])$). 正如所期望的, 这个关系式意味着 $P(1,0) = P(1,1)$.

对每一对 $0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1$, 设

$$\sigma(t,\lambda) := \frac{\partial P}{\partial \lambda}(t,\lambda) - h_j(\boldsymbol{G}(t,\lambda))\frac{\partial G_j}{\partial \lambda}(t,\lambda).$$

则假设 $\partial_i h_j = \partial_j h_i$ 在 Ω 中, $1 \leqslant i,j \leqslant N$, 以及关系式 $\dfrac{\partial}{\partial t}\left(\dfrac{\partial G_j}{\partial \lambda}\right) = \dfrac{\partial}{\partial \lambda}\left(\dfrac{\partial G_j}{\partial t}\right), 1 \leqslant$

$j \leqslant N$, 一起给出, 对每一对 $0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1$ 有

$$\begin{aligned}\frac{\partial \sigma}{\partial t}(t,\lambda) &= \frac{\partial}{\partial t}\left(\frac{\partial P}{\partial \lambda}\right)(t,\lambda) - \partial_i h_j(\boldsymbol{G}(t,\lambda))\frac{\partial G_i}{\partial t}(t,\lambda)\frac{\partial G_j}{\partial \lambda}(t,\lambda)\\ &\quad - h_j(\boldsymbol{G}(t,\lambda))\frac{\partial}{\partial t}\left(\frac{\partial G_j}{\partial \lambda}\right)(t,\lambda)\\ &= \frac{\partial}{\partial \lambda}\left(\frac{\partial P}{\partial t}\right)(t,\lambda) - \partial_j h_i(\boldsymbol{G}(t,\lambda))\frac{\partial G_j}{\partial \lambda}(t,\lambda)\frac{\partial G_i}{\partial t}(t,\lambda)\\ &\quad - h_j(\boldsymbol{G}(t,\lambda))\frac{\partial}{\partial \lambda}\left(\frac{\partial G_j}{\partial t}\right)(t,\lambda)\\ &= \frac{\partial}{\partial \lambda}\left(\frac{\partial P}{\partial t}\right)(t,\lambda) - \frac{\partial}{\partial \lambda}\left(h_i(\boldsymbol{G}(t,\lambda))\frac{\partial G_i}{\partial t}(t,\lambda)\right)\\ &= \frac{\partial}{\partial \lambda}\left(\frac{\partial P}{\partial t}(t,\lambda) - h_i(\boldsymbol{G}(t,\lambda))\frac{\partial G_i}{\partial t}(t,\lambda)\right) = 0,\end{aligned}$$

此因

$$\frac{\partial P}{\partial t}(t,\lambda) = h_i(\boldsymbol{G}(t,\lambda))\frac{\partial G_i}{\partial t}(t,\lambda), \quad 0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1.$$

注意到

$$\sigma(0,\lambda) = \frac{\partial P}{\partial \lambda}(0,\lambda) - h_j(\boldsymbol{G}(0,\lambda))\frac{\partial G_j}{\partial \lambda}(0,\lambda) = 0, \quad 0 \leqslant \lambda \leqslant 1$$

(由于 $P(0,\lambda) = P^0$ 与 $G_j(0,\lambda) = x_j^0, 1 \leqslant j \leqslant N$, 对所有 $0 \leqslant \lambda \leqslant 1$, 故 $\frac{\partial P}{\partial \lambda}(0,\lambda) = 0$ 与 $\frac{\partial G_j}{\partial \lambda}(0,\lambda) = 0$ 对所有 $0 \leqslant \lambda \leqslant 1$), 所以我们得到

$$0 = \sigma(1,\lambda) = \frac{\partial P}{\partial \lambda}(1,\lambda) - h_j(\boldsymbol{G}(1,\lambda))\frac{\partial G_j}{\partial \lambda}(1,\lambda) = \frac{\partial P}{\partial \lambda}(1,\lambda), \quad 0 \leqslant \lambda \leqslant 1$$

(此因 $G_j(1,\lambda) = x_j^1$ 对所有 $0 \leqslant \lambda \leqslant 1$).

(iii) 现在我们可以毫不含糊地定义一个函数 $p : \Omega \to \mathbb{R}$: 令

$$p(x^1) := P(1), \quad 对每个\ x^1 \in \Omega,$$

其中 $P \in \mathcal{C}^1[0,1]$ 是 (i) 中 Cauchy 问题的解, 其中的 $\gamma \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 是 Ω 中联结 x^0 与 x^1 的任意路径. 所以余下来的是要证明, $p \in \mathcal{C}^1(\Omega)$ 及 $\partial_i p = h_i$ 在 Ω 中, $1 \leqslant i \leqslant N$.

设给定一点 $x \in \Omega$ 和一个整数 $1 \leqslant i \leqslant N$. 则显然存在一点 $x^1 \in \Omega$, 一条在 Ω 中联结 x^0 和 x^1 的路径 $\gamma = (\gamma_i) \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$, 一个数 $\tau \in]0,1[$, 及一个包含 τ 的开区间 $I \subset [0,1]$, 使得

$$\gamma(t) = x + (t-\tau)\boldsymbol{e}_i, \quad 对\ t \in I,$$

其中 $\boldsymbol{e}_i$ 是 $\mathbb{R}^N$ 中第 i 个基向量.

设 $P \in \mathcal{C}^1[0,1]$ 表示 (i) 中相应于这条路径 γ 的 Cauchy 问题的解, 即 $p(\gamma(t)) = P(t), 0 \leqslant t \leqslant 1$. 则有

$$\begin{aligned} P(t) &= P(\tau) + (t-\tau)\frac{dP}{dt}(\tau) + o(t-\tau) \\ &= P(\tau) + (t-\tau)h_j(\gamma(\tau))\frac{d\gamma_j}{dt}(\tau) + o(t-\tau) \\ &= P(\tau) + (t-\tau)h_i(x) + o(t-\tau), \quad \text{对充分小的 } |t-\tau| \end{aligned}$$

(此处我们用到了 $\dfrac{d\gamma_j}{dt}(\tau) = \delta_{ij}$). 这就得到

$$p(x + (t-\tau)\boldsymbol{e}_i) = p(x) + (t-\tau)h_i(x) + o(t-\tau), \quad \text{对充分小的 } |t-\tau|,$$

此式说明 p 在 x 处具有由 $\partial_i p(x) = h_i(x)$ 给出的第 i 个偏导数 $\partial_i p$.

因为点 $x \in \Omega$ 和指标 $1 \leqslant i \leqslant N$ 是任意的, 而函数 h_i 在 Ω 中是 $\mathcal{C}^1$ 类的, 故函数 p 在 Ω 中是 $\mathcal{C}^2$ 类的且满足 $\partial_i p = h_i$ 在 $\mathcal{C}^1(\Omega)$ 中, $1 \leqslant i \leqslant N$.

(iv) 如果一个函数 $\pi \in \mathcal{C}^1(\Omega)$ 在 $\mathbb{R}^N$ 的一个连通开子集中对所有 $1 \leqslant i \leqslant N$ 满足 $\partial_i \pi = 0$, 则 π 是常数 (这个经典结果的证明将在更一般的意义下在定理 7.2-4 中给出). 所以 (iii) 中给出的函数 p 在模一个附加常数的意义下是唯一的. □

注 单连通性的假设是实质性的; 见习题 6.17-2.

作为定理 6.17-2 有用的补充, 我们现在证明, 任何函数 $p \in \mathcal{C}^2(\Omega)$ 均可以借助 Ω 中的路径积分用其梯度 $\mathbf{grad}\, p = (\partial_i p) \in \mathcal{C}^1(\Omega; \mathbb{R}^N)$ 表示出来; 而且在定理 6.17-2 的假设下, 这一路径积分为方程 $\mathbf{grad}\, p = \boldsymbol{h}$ 在 Ω 中提供了一个特解 p[86] (因此所有其他的解具有形式 $p + C, C$ 为常数).

定理 6.17-3 设 Ω 为 $\mathbb{R}^N$ 中的连通开子集, x^0 为 Ω 中一点.

(a) 给定任意函数 $p \in \mathcal{C}^2(\Omega)$, 设向量场 $\boldsymbol{h} = (h_i)_{i=1}^N \in \boldsymbol{\mathcal{C}}^1(\Omega)$ 由下式定义:

$$\boldsymbol{h} := \mathbf{grad}\, p.$$

则给定任一点 $x \in \Omega$ 及任一在 Ω 中联结 x^0 与 x 的路径 $\gamma_x \in \boldsymbol{\mathcal{C}}^\infty([0,1]; \mathbb{R}^N)$, 有

$$p(x) = p(x^0) + \int_{\gamma_x} \boldsymbol{h}(y) \cdot d\boldsymbol{y}, \quad \text{其中} \quad \int_{\gamma_x} \boldsymbol{h}(y) \cdot d\boldsymbol{y} := \int_{\gamma_x} h_i(y) dy_i.$$

(b) 假设 Ω 为单连通区域. 则给定任意满足

$$\mathbf{curl}\, \boldsymbol{h} = \mathbf{0} \quad \text{在 } \Omega \text{ 中}$$

的向量场 $\boldsymbol{h} \in \boldsymbol{\mathcal{C}}^1(\Omega)$ 及任意一点 $x \in \Omega$, 路径积分 $\displaystyle\int_{\gamma_x} \boldsymbol{h}(y) \cdot d\boldsymbol{y}$ 与联结 x^0 与 x 的路径 $\gamma_x \in \mathcal{C}^\infty([0,1]; \mathbb{R}^N)$ 的选取无关. 此外, 对任一这种路径, 由

$$p(x) := \int_{\gamma_x} \boldsymbol{h}(y) \cdot d\boldsymbol{y}, \quad x \in \Omega,$$

[86] 这个结果属于 Augustin-Louis Cauchy (1789—1857).

定义的函数 $p:\Omega\to\mathbb{R}$ 在 Ω 中是 $\mathcal{C}^2$ 类的, 而且是方程 $\mathbf{grad}\,p=\boldsymbol{h}$ 在 Ω 中的一个特解.

证明　给定任一点 $x\in\Omega$ 及任一联结 x^0 及 x 的路径 $\gamma_x=(\gamma_x^i)_{i=1}^N\in\mathcal{C}^\infty([0,1],\mathbb{R}^N)$, 则 $P(t)=p(\gamma_x(t)),0\leqslant t\leqslant 1$, 及定理 6.17-2 的证明 (i) 中的方程

$$\frac{dP}{dt}(t)=h_i(\gamma_x(t))\frac{d\gamma_x^i}{dt}(t),\quad 0\leqslant t\leqslant 1,$$

一并给出

$$P(1)=P(0)+\int_0^1 h_i(\gamma_x(t))\frac{d\gamma_x^i}{dt}(t)dt,$$

即

$$p(x)=p(x^0)+\int_{\gamma_x}h_i(y)dy_i,\quad \text{对任意 } x\in\Omega.$$

这就证明了 (a).

我们下面证明, 给定任意一个在 Ω 中满足 $\mathbf{curl}\,\boldsymbol{h}=\mathbf{0}$ 的向量场 $\boldsymbol{h}=(h_i)_{i=1}^N\in\boldsymbol{\mathcal{C}}^1(\Omega)$, 如果 Ω 是单连通的, 则积分 $\int_{\gamma_x}\boldsymbol{h}(y)\cdot d\boldsymbol{y}$ 与联结 x^0 与 x 的路径 $\gamma_x\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 无关. 为此, 设 $\gamma_x^0\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 及 $\gamma_x^1\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 是两个这种路径, 又设 $\boldsymbol{G}=(G_i)_{i=1}^N\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 是 Ω 中联结 γ_x^0 与 γ_x^1 的同伦. 则如定理 6.17-2 证明中 (ii) 所证, 对每个 $0\leqslant\lambda\leqslant 1$, 有

$$\int_{\boldsymbol{G}(\cdot,\lambda)}\boldsymbol{h}(y)\cdot d\boldsymbol{y}=\int_0^1\frac{\partial P}{\partial t}(t,\lambda)dt=P(1,\lambda)-P^0,$$

其中 $P(\cdot,\lambda)\in\boldsymbol{\mathcal{C}}^1[0,1]$ 表示 Cauchy 问题

$$\frac{\partial P}{\partial t}(t,\lambda)=h_i(\boldsymbol{G}(t,\lambda))\frac{\partial G_i}{\partial t}(t,\lambda),\quad 0\leqslant t\leqslant 1,\quad \text{及}\quad P(0,\lambda)=P^0$$

的唯一解.

在那里也证明了, 关系式 $\mathbf{curl}\,\boldsymbol{h}=\mathbf{0}$ 在 Ω 中, 意味着 $\dfrac{\partial P}{\partial\lambda}(1,\lambda)=0,0\leqslant\lambda\leqslant 1$, 故

$$\int_{\gamma_x^0}\boldsymbol{h}(\boldsymbol{y})\cdot d\boldsymbol{y}=P(1,0)-P^0=P(1,1)-P^0=\int_{\gamma_x^1}\boldsymbol{h}(\boldsymbol{y})\cdot d\boldsymbol{y}.$$

这就是所要证明的.

再用定理 6.17-2 证明中 (iii) 里同样的推导, 可以证明, 由

$$p:x\in\Omega\to p(x):=\int_{\gamma_x}\boldsymbol{h}(\boldsymbol{y})\cdot d\boldsymbol{y}$$

(上面已经证明, 这个定义是毫不含糊的) 定义的函数 $p:\Omega\to\mathbb{R}$ 在 Ω 中是可微的, 其偏导数为

$$\partial_i p(x)=h_i(x),\quad x\in\Omega,1\leqslant i\leqslant N.$$

这些关系式也说明函数 p 在 Ω 中是 $\mathcal{C}^2$ 类的. (b) 得证. □

J. L. Lions 引理的第三个应用是证明, 在实质性的较弱正则性假设下, 即向量场 $\boldsymbol{h}$ 的分量 $h_i, 1 \leqslant i \leqslant N$, 只是 $H^{-1}(\Omega)$ 中的分布时, Poincaré 引理仍然成立. 要注意, 经典的 Poincaré 引理以及 Δ 的次椭圆性在证明中仍起着关键的作用.

我们可以回忆一下, 仍为 J. L. Lions 引理和 Banach 闭值域定理的应用, 在定理 6.14-2 中已建立了对 $\boldsymbol{H}^{-1}(\Omega)$ 中向量场作为 $L^2(\Omega)$ 中数量函数梯度一种表征, 它与这里给出的是全然不同的.

定理 6.17-4 (弱 Poincaré 引理; 又称 $L^2(\Omega)$ 中的 Poincaré 引理[87]) 设 Ω 是 $\mathbb{R}^N$ 中的单连通区域, 又设给定一个向量场 $\boldsymbol{h} \in \boldsymbol{H}^{-1}(\Omega) := H^{-1}(\Omega; \mathbb{R}^N)$ 满足

$$\mathbf{curl}\, \boldsymbol{h} = \mathbf{0} \quad 在\ \boldsymbol{H}^{-2}(\Omega)\ 中.$$

则存在一个函数 $p \in L^2(\Omega)$ 使得

$$\mathbf{grad}\, p = \boldsymbol{h} \quad 在\ \boldsymbol{H}^{-1}(\Omega)\ 中.$$

此外, 方程 $\partial_i \tilde{p} = h_i$ 在 $H^{-1}(\Omega)$ 中, $1 \leqslant i \leqslant N$, 的任何另一解 $\tilde{p} \in L^2(\Omega)$ 具有形式 $\tilde{p} = p + C$, 其中 C 为常数.

证明　首先回忆一下, 梯度算子 $\mathbf{grad} : \mathcal{D}'(\Omega) \to \boldsymbol{\mathcal{D}}'(\Omega)$ 定义为

$$(\mathbf{grad}\, v)_i := \partial_i v, \quad 1 \leqslant i \leqslant N,\ 对每个\ v \in \mathcal{D}'(\Omega),$$

散度算子 $\mathrm{div} : \boldsymbol{\mathcal{D}}'(\Omega) \to \mathcal{D}'(\Omega)$ 定义为

$$\mathrm{div}\, \boldsymbol{v} := \sum_{i=1}^{N} \partial_i v_i, \quad 对每个\ \boldsymbol{v} = (v_i)_{i=1}^N \in \boldsymbol{\mathcal{D}}'(\Omega),$$

向量 Laplace 算子 $\Delta : \boldsymbol{\mathcal{D}}'(\Omega) \to \boldsymbol{\mathcal{D}}'(\Omega)$ 定义为

$$(\Delta \boldsymbol{v})_i := \Delta v_i, \quad 1 \leqslant i \leqslant N,\ 对每个\ \boldsymbol{v} = (v_i)_{i=1}^N \in \boldsymbol{\mathcal{D}}'(\Omega),$$

而旋度算子 $\mathbf{curl} : \boldsymbol{\mathcal{D}}'(\Omega) \to \mathcal{D}'(\Omega; \mathbb{R}^{\frac{N(N-1)}{2}})$ 对任意整数 $N \geqslant 2$ 定义为

$$(\mathbf{curl}\, \boldsymbol{v})_{ij} := (\partial_i v_j - \partial_j v_i), \quad 1 \leqslant i < j \leqslant N,\ 对每个\ \boldsymbol{v} = (v_i) \in \boldsymbol{\mathcal{D}}'(\Omega).$$

---

[87] 这个结果属于:

P. G. Ciarlet; P. Ciarlet, Jr. [2005]: Another approach to linearized elasticity and a new proof of Korn's inequality. Mathematical Models and Methods in Applied Sciences **15**, 259–271.

这里给出的较简单的证明属于:

S. Kesavan [2005]: On Poincaré's and J. L. Lions' lemmas. Comptes Rendus de l'Académie des Sciences de Paris, Série I, **340**, 27–30.

Poincaré 引理在即使较弱的分布意义下也成立, 这一结果获证较晚, 见:

S. Mardare [2008]: On Poincaré and De Rham's theorems. Revue Roumaine de Mathématiques Pures et Appliquées **53**, 523–541.

我们现在要证明, 如果 $\boldsymbol{h} \in \boldsymbol{H}^{-1}(\Omega)$ 满足 $\mathbf{curl}\ \boldsymbol{h} = \mathbf{0}$ 在 $\boldsymbol{H}^{-2}(\Omega)$ 中, 则存在 $p \in L^2(\Omega)$ 使得 $\boldsymbol{h} = \mathbf{grad}\, p$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中. 为此目的, 我们分两步进行.

(i) 由于定理 6.14-3 (其证明依赖 J. L. Lions 引理, 用定理 6.14-1 的方法) 在右端 $\boldsymbol{h}$ 属于 $\boldsymbol{H}^{-1}(\Omega)$ 而不是 $\boldsymbol{L}^2(\Omega)$ 时同样成立, 故存在一个向量场 $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ 和一个函数 $\lambda \in L^2(\Omega)$ 使得

$$-\Delta \boldsymbol{u} + \mathbf{grad}\, \lambda = \boldsymbol{h} \quad 在\ \boldsymbol{H}^{-1}(\Omega)\ 中,$$
$$\operatorname{div} \boldsymbol{u} = 0 \quad 在\ L^2(\Omega)\ 中.$$

注意, 在这个阶段, Ω 是单连通的及 $\mathbf{curl}\, \boldsymbol{h} = \mathbf{0}$ 在 $\boldsymbol{H}^{-2}(\Omega)$ 中这些假设并不需要.

(ii) 假设 $\mathbf{curl}\, \boldsymbol{h} = \mathbf{0}$ 在 $\boldsymbol{H}^{-2}(\Omega)$ 中, 连同关系式

$$\mathbf{curl}\,\mathbf{grad}\, \pi = \mathbf{0} \quad 在\ \boldsymbol{\mathcal{D}}'(\Omega)\ 中, \quad 对任意\ \pi \in \mathcal{D}'(\Omega),$$

意味着

$$\Delta(\mathbf{curl}\, \boldsymbol{u}) = \mathbf{curl}(\Delta \boldsymbol{u}) = \mathbf{curl}\,\mathbf{grad}\lambda - \mathbf{curl}\, \boldsymbol{h} = \mathbf{0}.$$

因 $\mathbf{curl}\, \boldsymbol{u} \in \boldsymbol{L}^2(\Omega) \subset \boldsymbol{L}^1_{\mathrm{loc}}(\Omega)$, Δ 的次椭圆性 (定理 6.4-2) 说明

$$\mathbf{curl}\, \boldsymbol{u} \in \mathcal{C}^\infty(\Omega),$$

故 $(\partial_j u_i - \partial_i u_j) \in \mathcal{C}^\infty(\Omega)$ 对所有 $1 \leqslant i, j \leqslant N$. 所以

$$\sum_{j=1}^N \partial_j(\partial_j u_i - \partial_i u_j) = \Delta u_i - \partial_i(\operatorname{div} u) = \Delta u_i \in \mathcal{C}^\infty(\Omega), \quad 1 \leqslant i \leqslant N,$$

此因 $\operatorname{div} \boldsymbol{u} = 0$.

因为 $\Delta \boldsymbol{u} \in \mathcal{C}^\infty(\Omega)$ 及 $\mathbf{curl}\, \Delta \boldsymbol{u} = \mathbf{0}$ 在 Ω 中, 而且 Ω 是单连通的, 经典 Poincaré 引理 (定理 6.17-2) 适用, 即存在一个函数 $\tilde{p} \in \mathcal{C}^\infty(\Omega) \subset L^1_{\mathrm{loc}}(\Omega) \subset \mathcal{D}'(\Omega)$ 使得

$$\mathbf{grad}\, \tilde{p} = \Delta \boldsymbol{u} = \mathbf{grad}\, \lambda - \boldsymbol{h} \quad 在\ \boldsymbol{H}^{-1}(\Omega)\ 中.$$

又因为分布

$$p := \lambda - \tilde{p} \in L^1_{\mathrm{loc}}(\Omega)$$

满足

$$\mathbf{grad}\, p = \mathbf{grad}\, \lambda - \mathbf{grad}\, \tilde{p} = \boldsymbol{h} \in \boldsymbol{H}^{-1}(\Omega),$$

J. L. Lions 引理 (定理 6.11-4) 说明, p 实际上是 $L^2(\Omega)$ 中的函数.

方程 $\mathbf{grad}\, p = \boldsymbol{h}$ 在 $\boldsymbol{H}^{-1}(\Omega)$ 中的解 $p \in L^2(\Omega)$ 在相差一个常数意义下的唯一性, 可如定理 6.14-2 的证明中那样予以确立. □

注 还是利用 Δ 的次椭圆性可以证明向量场 $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ 也在空间 $\mathcal{C}^\infty(\Omega)$ 中, 此因 $\Delta \boldsymbol{u} \in \mathcal{C}^\infty(\Omega)$. 然而这一性质在上述证明中并未用到.

J. L. Lions 引理连同 Δ 的次椭圆性在弱 Poincaré 引理的证明中起着关键的作用. 值得注意的是, 弱 Poincaré 引理反过来又为 J. L. Lions 引理提供了一个非常简单的证明 (习题 6.17-3).

习题

6.17-1 设 Ω 是 $\mathbb{R}^N$ 的开子集.

(1) 设给定一个向量场 $\boldsymbol{h} \in \boldsymbol{\mathcal{C}}^1(\Omega)$ 及一点 $x_0 \in \Omega$, 使得对任一点 $x \in \Omega$, 曲线积分 $\int_{\gamma(x)} \boldsymbol{h}(y) \cdot d\boldsymbol{y}$ 与 Ω 中联结 x_0 和 x 的路径 $\gamma(x)$ 无关. 证明 $\mathbf{curl}\, \boldsymbol{h} = \mathbf{0}$ 在 Ω 中.

(2) 设给定一个向量场 $\boldsymbol{h} \in \boldsymbol{\mathcal{C}}^1(\Omega)$. 证明, 存在 $p \in \mathcal{C}^2(\Omega)$ 使得 $\mathbf{grad}\, \boldsymbol{p} = \boldsymbol{h}$ 在 Ω 中成立的充分必要条件是, 曲线积分 $\int_{\gamma} \boldsymbol{h}(y) \cdot d\boldsymbol{y}$ 与在 Ω 中联结任意两不同点的路径 γ 无关.

6.17-2 本题在 Ω 不是单连通的情况下, 提供了一个既针对经典的也针对弱 Poincaré 引理 (定理 6.17-2 及定理 6.17-4) 的反例.

(1) 设

$$\Omega := \{(x_1, x_2) \in \mathbb{R}^2; 1 < x_1^2 + x_2^2 < 2\},$$
$$h_1(x_1, x_2) := \frac{x_2}{x_1^2 + x_2^2} \quad \text{及} \quad h_2(x_1, x_2) := -\frac{x_1}{x_1^2 + x_2^2} \quad \text{对 } (x_1, x_2) \in \Omega,$$

它们满足 $h_1, h_2 \in \mathcal{C}^\infty(\Omega)$ 且 $\partial_1 h_2 - \partial_2 h_1 = 0$ 在 Ω 中.

证明, 不存在任何函数 $p \in \mathcal{C}^2(\Omega)$ 满足 $\partial_i p = h_i$ 在 Ω 中, $i = 1, 2$.

提示: 令 $\tilde{\Omega} := \Omega - \gamma$, 其中 $\gamma := \{(x_1, 0) \in \mathbb{R}^2; -2 < x_1 < -1\}$. 然后计算方程 $\partial_i p = h_i$ 在 $\tilde{\Omega}$ 中, $i = 1, 2$, 的显式一般解, 并证明 $\lim_{x_2 \to 0^+} p(x_1, x_2) \neq \lim_{x_2 \to 0^-} p(x_1, x_2)$ 对所有 $-2 < x_1 < -1$.

(2) 对任意维数 $N \geqslant 3$, 构造一个类似的反例.

6.17-3 (1) 设 Ω 是 $\mathbb{R}^N$ 中的单连通区域, 而 $v \in \mathcal{D}'(\Omega)$ 是满足 $\mathbf{grad}\, v \in \boldsymbol{H}^{-1}(\Omega)$ 的一个分布. 假定弱 Poincaré 引理 (定理 6.17-4) 成立, 利用验证 $v \in L^2(\Omega)$ 给出 J. L. Lions 引理 (定理 6.11-4) 一个只有两行的简单证明.

(2) 假定 J. L. Lions 引理对 $\mathbb{R}^N$ 中的单连通区域成立, 证明它对 $\mathbb{R}^N$ 中的任何区域均成立.

6.18 Poincaré 引理的应用: 经典的和弱 Saint-Venant 引理; Cesàro-Volterra 路径积分公式

本节是 6.17 节的 "矩阵类似", 矩阵对称化梯度算子

$$\boldsymbol{v} : \mathcal{D}'(\Omega; \mathbb{R}^N) \to \boldsymbol{\nabla}_s(\boldsymbol{v}) := \frac{1}{2}(\boldsymbol{\nabla} \boldsymbol{v}^{\mathrm{T}} + \boldsymbol{\nabla} \boldsymbol{v}) \in \mathcal{D}'(\Omega; \mathbb{S}^N)$$

起着向量梯度算子

$$\mathbf{grad} : p \in \mathcal{D}'(\Omega) \to \mathbf{grad}\, \boldsymbol{p} \in \mathcal{D}'(\Omega; \mathbb{R}^N)$$

的作用. 这就是为什么我们下面的论述沿着 6.17 节同样路线进行的缘由. 要注意的是, 下面的定理 6.18-1 及定理 6.18-3 的关键均取决于 Poinceré 引理, 分别取决于其经典形式与弱形式.

在整节中, 仍用关于重复指标的求和约定. 给定 $\mathbb{R}^N$ 中一开子集 Ω, 考虑由下式定义的从空间 $\mathcal{C}^3(\Omega) := \mathcal{C}^3(\Omega; \mathbb{R}^N)$ 到空间 $\mathcal{C}^2(\Omega; \mathbb{S}^N)$ 的线性算子:

$$\boldsymbol{v} = (v_i) \in \mathcal{C}^3(\Omega) \to \boldsymbol{\nabla}_s \boldsymbol{v} = \left(\frac{1}{2}(\partial_j v_i + \partial_i v_j)\right) \in \mathcal{C}^2(\Omega; \mathbb{S}^N).$$

回忆一下, 矩阵场 $\boldsymbol{\nabla}_s \boldsymbol{v}$ 曾在定理 6.15-1 的基本 Korn 不等式中出现过 (在那里它被表示为 $\boldsymbol{e}(\boldsymbol{v}) = (e_{ij}(\boldsymbol{v}))$).

一个自然的问题就此出现, 就是这个线性算子是否可逆, 即给定一个矩阵场 $\boldsymbol{e} = (e_{ij}) \in \mathcal{C}^2(\Omega; \mathbb{S}^N)$, 是否存在一个向量场 $\boldsymbol{v} = (v_i) \in \mathcal{C}^3(\Omega; \mathbb{R}^N)$ 使得

$$\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij} \text{ 在 } \Omega \text{ 中}, \quad 1 \leqslant i, j \leqslant N,$$

或等价地使得

$$\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e} \quad \text{在 } \Omega \text{ 中}.$$

如果是肯定答案, 那立即可得, 函数 $e_{ij} = e_{ji} \in \mathcal{C}^2(\Omega)$ 必须满足 Saint-Venant 相容性关系[88]:

$$\partial_{\ell j} e_{ik} + \partial_{ki} e_{j\ell} - \partial_{\ell i} e_{jk} - \partial_{kj} e_{i\ell} = 0 \text{ 在 } \mathcal{C}(\Omega) \text{ 中}, \quad 1 \leqslant i, j, k, l \leqslant N.$$

注 当 $N = 3$ 时, Saint-Venant 相容性关系可以精简为一个单一的矩阵方程; 见习题 6.18-4.

值得注意的是, 如果开集 Ω 是单连通的, 那么这些必要条件也就变成充分的了.

定理 6.18-1 (经典 Saint-Venant 引理; 也称 $\mathcal{C}^3(\Omega)$ 中的 Saint-Venant 引理) 设 Ω 是 $\mathbb{R}^N$ 的单连通开子集, 另设给定函数 $e_{ij} = e_{ji} \in \mathcal{C}^2(\Omega), 1 \leqslant i, j \leqslant N$, 满足 Saint-Venant 相容性关系

$$\partial_{\ell j} e_{ik} + \partial_{ki} e_{j\ell} - \partial_{\ell i} e_{jk} - \partial_{kj} e_{i\ell} = 0 \text{ 在 } \Omega \text{ 中}, \quad 1 \leqslant i, j, k, \ell \leqslant N.$$

则存在向量场 $\boldsymbol{v} = (v_i) \in \mathcal{C}^3(\Omega; \mathbb{R}^N)$ 使得

$$\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij} \text{ 在 } \Omega \text{ 中}, \quad 1 \leqslant i, j \leqslant N.$$

[88] 冠名源自 Adhémar-Jean-Claude Barré de Saint-Venant (1797—1886), 他于 1864 年发表了这些关系.

此外, 方程

$$\frac{1}{2}(\partial_j \tilde{v}_i + \partial_i \tilde{v}_j) = e_{ij} \text{ 在 } \Omega \text{ 中}, \quad 1 \leqslant i, j \leqslant N,$$

的任一其他解 $\tilde{\boldsymbol{v}} = (\tilde{\boldsymbol{v}}_i) \in \mathcal{C}^3(\Omega; \mathbb{R}^N)$ 都具有如下形式:

$$\tilde{\boldsymbol{v}}(x) = \boldsymbol{v}(x) + \boldsymbol{B}x + \boldsymbol{c}, \quad x \in \Omega,$$

其中 $\boldsymbol{B}$ 为某 $N \times N$ 反对称矩阵, $\boldsymbol{c} \in \mathbb{R}^N$ 为某向量.

证明 对于本证明给出的各种关系式, 我们都默认它们对出现在其中的拉丁下标的所有值 $1, 2, \cdots, N$ 均成立. Saint-Venant 相容性关系可以等价地重写为

$$\partial_\ell h_{ijk} = \partial_k h_{ij\ell} \text{ 在 } \mathcal{C}(\Omega) \text{ 中}, \quad \text{其中} \quad h_{ijk} := \partial_j e_{ik} - \partial_i e_{jk} \in \mathcal{C}^1(\Omega).$$

所以经典 Poincaré 引理 (定理 6.17-2) 说明, 存在函数 $p_{ij} \in \mathcal{C}^2(\Omega)$, 其在相差一个常数的意义下是唯一的, 使得

$$\partial_k p_{ij} = h_{ijk} = \partial_j e_{ik} - \partial_i e_{jk} \quad \text{在 } \mathcal{C}^1(\Omega) \text{ 中}.$$

此外, 由 $\partial_k p_{ij} = -\partial_k p_{ji}$ 在 $\mathcal{C}^1(\Omega)$ 中, 我们就有了选取函数 p_{ij} 的自由度, 使其满足 $p_{ij} + p_{ji} = 0$ 在 $\mathcal{C}^2(\Omega)$ 中.

注意到函数 $q_{ij} := (e_{ij} + p_{ij}) \in \mathcal{C}^2(\Omega)$ 满足

$$\begin{aligned}\partial_k q_{ij} &= \partial_k e_{ij} + \partial_k p_{ij} = \partial_k e_{ij} + \partial_j e_{ik} - \partial_i e_{jk} \\ &= \partial_j e_{ik} + \partial_j p_{ik} = \partial_j q_{ik} \quad \text{在 } \mathcal{C}^1(\Omega) \text{ 中},\end{aligned}$$

我们又可以借助经典 Poincaré 引理断定, 存在函数 $v_i \in \mathcal{C}^3(\Omega)$, 在相差一个常数的意义下是唯一的, 使得

$$\partial_j v_i = q_{ij} = e_{ij} + p_{ij} \quad \text{在 } \Omega \text{ 中}.$$

所以

$$\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij} + \frac{1}{2}(p_{ij} + p_{ji}) = e_{ij} \quad \text{在 } \Omega \text{ 中},$$

这正是所需证明的. 所有其他的解具有定理中指出的形式这一结论, 可如定理 6.15-2 的证明中那样予以确立. □

注 单连通性的假设是实质性的; 见习题 6.18-2.

作为定理 6.18-1 的一个有用的补充, 我们现在证明, 在定理 6.18-1 的假设下, 任意向量场 $\boldsymbol{v} \in \mathcal{C}^3(\Omega; \mathbb{R}^N)$ 可以通过在 Ω 中的路径积分, 用其对称化梯度场 $\boldsymbol{\nabla}_s \boldsymbol{v} = \left(\frac{1}{2}(\partial_j v_i + \partial_i v_j)\right) \in \mathcal{C}^2(\Omega; \mathbb{S}^N)$ 的分量表示出来. 这同一个路径积分也为方程 $\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e}$ 在 Ω 中提供了一个特解 (因此所有其他的解均可由这个特解加上一个形如 $x \in \Omega \to \boldsymbol{B}x + \boldsymbol{c}$ 的向量场得到, 其中 $\boldsymbol{B}$ 是 $N \times N$ 反对称矩阵, 而 $\boldsymbol{c} \in \mathbb{R}^N$).

定理 6.18-2 (Cesàro-Volterra 路径积分公式[89]) 设 Ω 是 $\mathbb{R}^N$ 的连通开子集, x^0 是 Ω 中的一点.

(a) 给定任意向量场 $(v_i) \in \mathcal{C}^3(\Omega;\mathbb{R}^N)$, 由下式定义对称张量场 $(e_{ij}) \in \mathcal{C}^2(\Omega;\mathbb{S}^N)$:

$$e_{ij} := \frac{1}{2}(\partial_j v_i + \partial_i v_j), \quad 1 \leqslant i,j \leqslant N.$$

则给定任意点 $x = (x_k) \in \Omega$ 及 Ω 中联结 x^0 与 x 的任意路径 $\gamma_x \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$, 向量场的分量由 Cesàro-Volterra 路径积分公式给出, 即

$$v_i(x) = v_i^0 + p_{ik}^0(x_k - x_k^0) + \int_{\gamma_x} \{e_{ij}(y) + (\partial_k e_{ij}(y) - \partial_i e_{kj}(y))\}(x_k - y_k)dy_j, \quad 1 \leqslant i \leqslant N,$$

其中

$$v_i^0 := v_i(x^0) \quad 及 \quad p_{ik}^0 := \frac{1}{2}(\partial_k v_i(x^0) - \partial_i v_k(x^0)).$$

(b) 假设 Ω 是单连通的. 给定任意对称张量场 $(e_{ij}) \in \mathcal{C}^2(\Omega;\mathbb{S}^N)$, 其分量满足 Saint-Venant 相容性关系

$$\partial_{\ell j} e_{ik} + \partial_{ki} e_{j\ell} - \partial_{\ell i} e_{jk} - \partial_{kj} e_{i\ell} = 0 \text{ 在 } \Omega \text{ 中}, \quad 1 \leqslant i,j,k,\ell \leqslant N,$$

又给定任意点 $x \in \Omega$, 则每一个路径积分 $\int_{\gamma_x} (e_{ij}(y) + (\partial_k e_{ij}(y) - \partial_i e_{kj}(y))(x_k - y_k)dy_j, 1 \leqslant i \leqslant N$, 均与所选取的联结 x^0 与 x 的路径 $\gamma_x \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 无关. 此外, 对每个 $x = (x_k) \in \Omega$, 由

$$v_i(x) := \int_{\gamma_x} \{e_{ij}(y) + (\partial_k e_{ij}(y) - \partial_i e_{kj}(y))\}(x_k - y_k)dy_j, \quad 1 \leqslant i \leqslant N$$

定义的向量场 $(v_i) : \Omega \to \mathbb{R}^N$ 在 Ω 中是 $\mathcal{C}^2$ 类的, 而且是方程

$$\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij} \text{ 在 } \Omega \text{ 中}, \quad 1 \leqslant i,j \leqslant N,$$

的一个特解.

证明 给定任意的 $x \in \Omega$ 及任意联结 x^0 与 x 的路径 $\gamma_x = (\gamma_x^i)_{i=1}^N \in \mathcal{C}^\infty([0,1];\mathbb{R}^N)$, 我们有

$$\begin{aligned} v_i(x) &= v_i(x^0) + \int_0^1 \frac{d}{dt}[v_i(\gamma_x(t))]dt = v_i(x^0) + \int_0^1 \partial_j v_i(\gamma_x(t)) \frac{d\gamma_x^j}{dt}(t)dt \\ &= v_i(x^0) + \int_{\gamma_x} \partial_j v_i(y)dy_j = v_i(x^0) + \int_{\gamma_x} e_{ij}(y)dy_j + \int_{\gamma_x} p_{ij}(y)dy_j, \end{aligned}$$

[89] 应归之于:

E. Cesàro [1906]: Sulle formole del Volterra, fondamentali nella teoria delle distorsioni elastiche. Rendiconti Napoli **12**, 311–321.

V. Volterra [1907]: Sur l'équilibre des corps élastiques multiplement connexes. Annales de l'Ecole Normale **24**, 401–517.

其中函数 $p_{ij} \in \mathcal{C}^2(\Omega)$ 定义为

$$p_{ij} := \frac{1}{2}(\partial_j v_i - \partial_i v_j) \quad \text{在 } \Omega \text{ 中}.$$

注意到

$$\begin{aligned}
\int_{\gamma_x} p_{ij}(y)dy_j &= \int_0^1 p_{ij}(\gamma_x(t))\frac{d\gamma_x^j}{dt}(t)dt \\
&= -\int_0^1 \left(\frac{d}{dt}[p_{ij}(\gamma_x(t))]\right)\gamma_x^j(t)df + p_{ij}(x)\gamma_x^j(1) - p_{ij}(x^0)\gamma_x^j(0) \\
&= -\int_{\gamma_x} \partial_j p_{ik}(y)y_k dy_j + p_{ik}(x)x_k - p_{ik}^0 x_k^0,
\end{aligned}$$

以及

$$\int_{\gamma_x} x_k \partial_j p_{ik}(y)dy_j = x_k \int_0^1 \frac{d}{dt}[p_{ik}(\gamma_x(t))]dt = p_{ik}(x)x_k - p_{ik}^0 x_k,$$

我们就得到

$$\int_{\gamma_x} p_{ij}(y)dy_j = p_{ik}^0(x_k - x_k^0) + \int_{\gamma_x} \partial_j p_{ik}(y)(x_k - y_k)dy_j.$$

这样, 从关系式

$$\partial_j p_{ik} = \partial_k e_{ij} - \partial_i e_{kj}$$

就得到 Cesàro-Volterra 公式. 这就证明了 (a).

(b) 的证明类似于定理 6.17-3 中 (b) 的相应内容. 为此, 留给读者作为习题 (习题 6.18-5). □

注 上述借助于 Cesàro-Volterra 路径积分公式, 利用其对称化梯度来表示一个向量场的表达式, 与 Calderón-Zygmund 奇异积分的 (精细) 理论相结合, 可以给出 Korn 不等式一个直接的证明[90).

注 当 $N = 3$ 时, 构成 Cesàro-Volterra 路径积分公式的三个关系式可方便地精简为单一的向量方程; 参见习题 6.18-4.

利用 Poincaré 引理的弱形式, 当然因此最终还是要再次用到 J. L. Lions 引理, 我们现在证明, 在一个本质上较弱的正则性假设下, 即 $e_{ij}, 1 \leqslant i, j \leqslant N$, 只是 $L^2(\Omega)$ 中的函数, Saint-Venant 引理仍然成立.

有趣的是, 这个 Saint-Venant 引理的 "弱形式" 也为 Korn 不等式提供了一个新的证明 (定理 6.18-5).

[90)] 这种方法应归之于:

P. P. Mosolov; V. P. Mjasnikov [1971]: A proof of Korn's inequality. Soviet Mathematics Doklady **12**, 1618–1622.

定理 6.18-3 (弱 Saint-Venant 引理; 也称 $H^1(\Omega)$ 中的 Saint-Venant 引理[91]) 设 Ω 是 $\mathbb{R}^N$ 中的单连通区域, 又设 $\boldsymbol{e} = (e_{ij}) \in \mathbb{L}^2(\Omega)$ 是对称矩阵场, 满足 Saint-Venant 相容性关系:

$$\partial_{\ell j} e_{ik} + \partial_{ki} e_{j\ell} - \partial_{\ell i} e_{jk} - \partial_{kj} e_{i\ell} = 0 \text{ 在 } H^{-2}(\Omega) \text{ 中}, \quad 1 \leqslant i, j, k, \ell \leqslant N.$$

则存在一个向量场 $\boldsymbol{v} = (v_i)_{i=1}^N \in \boldsymbol{H}^1(\Omega)$ 使得

$$\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij} \text{ 在 } L^2(\Omega) \text{ 中}, \quad 1 \leqslant i, j \leqslant N.$$

此外, 方程 $e_{ij} = \dfrac{1}{2}(\partial_j \tilde{v}_i + \partial_i \tilde{v}_j), 1 \leqslant i, j \leqslant N$, 所有的其他解 $\tilde{\boldsymbol{v}} = (\tilde{v}_i)_{i=1}^N \in \boldsymbol{H}^1(\Omega)$ 具有以下形式

$$\tilde{\boldsymbol{v}}(x) = \boldsymbol{v}(x) + \boldsymbol{B}x + \boldsymbol{c}, \quad \text{对几乎所有 } x \in \Omega,$$

其中 $\boldsymbol{B}$ 是某 $N \times N$ 反对称矩阵而 $\boldsymbol{c} \in \mathbb{R}^N$ 为某向量.

证明 这个定理的证明类似于定理 6.18-1 中的, 只是现在用的是 Poincaré 引理的弱形式 (定理 6.17-4). 首先证明存在函数 $p_{ij} \in L^2(\Omega)$, 在相差一个常数的意义下是唯一的, 满足

$$\partial_k p_{ij} = h_{ijk} = \partial_j e_{ik} - \partial_i e_{jk} \quad \text{在 } H^{-1}(\Omega) \text{ 中};$$

其次证明存在函数 $v_i \in H^1(\Omega)$, 也是在相差一个常数意义下是唯一的, 满足 $\partial_j v_i = q_{ij} = e_{ij} + p_{ij}$ 在 $L^2(\Omega)$ 中.

所以有

$$\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij} + \frac{1}{2}(p_{ij} + p_{ji}) = e_{ij} \quad \text{在 } L^2(\Omega) \text{ 中},$$

这就是所要证明的. 由定理 6.15-2 可得, 所有其他的解均具有定理中指出的形式. □

注 对于一张量 $\boldsymbol{e} \in \mathbb{L}^2(\Omega)$ 可表示为如下形式: $\boldsymbol{e} = \dfrac{1}{2}(\boldsymbol{\nabla v}^{\mathrm{T}} + \boldsymbol{\nabla v})$ 对某个 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$, 一个不同的充分必要条件将在下节 (定理 6.19-6) 中给出. 这个条件断言, 张量 $\boldsymbol{e}$ 应该在由所有在 Ω 中满足 $\operatorname{div} \boldsymbol{\sigma} = 0$ 的对称张量 $\boldsymbol{\sigma} \in \mathbb{H}_0^1(\Omega)$ 张成空间 (在 $\mathbb{H}_0^1(\Omega)$ 里) 的正交补中; 此外, 并不要求开集 Ω 是单连通的.

---

[91] 这个结果属于:

P. G. Ciarlet; P. Ciarlet, Jr. [2005]: Another approach to linearized elasticity and a new proof of Korn's inequality. Mathematical Models and Methods in Applied Sciences **15**, 259–271.

其各种扩展推广可参阅:

G. Geymonat; F. Krasucki [2005]: Some remarks on the compatibility conditions in elasticity. Accademia Nazionale delle Scienze detta dei XL. Rendiconti. Serie V. Memorie di Matematica e Applicazioni. Parte I, **29**, 175–181.

G. Geymonat; F. Krasucki [2006]: Beltrami's solutions of general equilibrium equations in continuum mechanics. Comptes Rendus de l'Académie des Sciences de Paris, Série 1, **342**, 359–363.

C. Amrouche; P.G. Ciarlet; L. Gratie; S. Kesavan [2006]: On the characterization of matrix fields as linearized strain tensor fields. Journal de Mathématiques Pures et Appliquées **86**, 116–132.

设一对称矩阵场 $\boldsymbol{e} = (e_{ij}) \in \mathbb{L}^2(\Omega)$ 满足

$$\partial_{\ell j} e_{ik} + \partial_{ki} e_{j\ell} - \partial_{\ell i} e_{jk} - \partial_{kj} e_{i\ell} = 0 \quad \text{在 } H^{-2}(\Omega) \text{ 中}, \quad 1 \leqslant i, j, k, \ell \leqslant N,$$

即 Saint-Venant 相容性关系的弱形式. 由定理 6.18-3, 存在唯一的等价类 $\dot{\boldsymbol{v}} \in \dot{\boldsymbol{H}}^1(\Omega) = \boldsymbol{H}^1(\Omega)/\mathbf{Ker}\ \boldsymbol{\nabla}_s$ 使得 $\boldsymbol{e} = \boldsymbol{\nabla}_s \dot{\boldsymbol{v}}$ 在 $\mathbb{L}^2(\Omega)$ 中, 其中 (定理 6.15-2)

$$\mathbf{Ker}\ \boldsymbol{\nabla}_s = \{\boldsymbol{v} \in \boldsymbol{H}^1(\Omega);\ \text{存在 } \boldsymbol{B} \in \mathbb{A}^N \text{ 与 } \boldsymbol{c} \in \mathbb{R}^N \text{ 使得 } \boldsymbol{v}(x) = \boldsymbol{B}x + \boldsymbol{c} \text{ 对几乎所有 } x \in \Omega\}.$$

我们就在证明, 以上述方式定义的映射 $\Xi : \boldsymbol{e} \in \mathbb{L}^2(\Omega) \to \dot{\boldsymbol{v}} \in \dot{\boldsymbol{H}}^1(\Omega)$ 具有值得关注的性质.

定理 6.18-4 设 Ω 是 $\mathbb{R}^N$ 中的单连通区域. 定义空间

$$\begin{aligned}\mathbb{E}(\Omega) := \{\boldsymbol{e} = (e_{ij}) \in \mathbb{L}^2(\Omega); \partial_{\ell j} e_{ik} + \partial_{ki} e_{j\ell} - \partial_{\ell i} e_{jk} - \partial_{kj} e_{i\ell} = 0 \\ \text{在 } H^{-2}(\Omega) \text{ 中},\ 1 \leqslant i, j, k, \ell \leqslant N\},\end{aligned}$$

又设

$$\Xi : \mathbb{E}(\Omega) \to \dot{\boldsymbol{H}}^1(\Omega)$$

是对每个 $\boldsymbol{e} \in \mathbb{E}(\Omega)$ 由 $\Xi(\boldsymbol{e}) := \dot{\boldsymbol{v}}$ 定义的线性映射, 其中 $\dot{\boldsymbol{v}}$ 是在商空间 $\dot{\boldsymbol{H}}^1(\Omega)$ 中满足

$$\boldsymbol{\nabla}_s \dot{\boldsymbol{v}} = \boldsymbol{e} \quad \text{在 } \mathbb{L}^2(\Omega) \text{ 中}$$

的唯一元素 (定理 6.18-3). 则

$$\Xi \in \mathcal{L}(\mathbb{E}(\Omega); \dot{\boldsymbol{H}}^1(\Omega)),\ \Xi \text{ 是双射, 且 } \Xi^{-1} \in \mathcal{L}(\dot{\boldsymbol{H}}^1(\Omega); \mathbb{E}(\Omega)).$$

证明 显然, $\mathbb{E}(\Omega)$ 作为 $\mathbb{L}^2(\Omega)$ 的闭子空间, 是 Hilbert 空间. 映射 Ξ 是单射, 此因 $\Xi(e) = \dot{\mathbf{0}}$ 意味着 $\boldsymbol{e} = \boldsymbol{\nabla}_s \dot{\mathbf{0}} = \mathbf{0}$; 它也是满射, 此因给定任意的 $\dot{\boldsymbol{v}} \in \dot{\boldsymbol{H}}^1(\Omega)$, 矩阵场 $(e_{ij}) := \boldsymbol{\nabla}_s \dot{\boldsymbol{v}} \in \mathbb{L}^2(\Omega)$ 必须满足 $\partial_{\ell j} e_{ik} + \partial_{ki} e_{j\ell} - \partial_{\ell i} e_{jk} - \partial_{kj} e_{i\ell} = 0$ 在 $H^{-2}(\Omega)$ 中.

最后, 逆映射

$$\Xi^{-1} : \dot{\boldsymbol{v}} \in \dot{\boldsymbol{H}}^1(\Omega) \to \boldsymbol{\nabla}_s \dot{\boldsymbol{v}} \in \mathbb{E}(\Omega)$$

是连续的, 这是因为, 显然存在常数 c 使得

$$\|\boldsymbol{\nabla}_s \dot{\boldsymbol{v}}\|_{0,\Omega} = \|\boldsymbol{\nabla}_s (\boldsymbol{v} + \boldsymbol{r})\|_{0,\Omega} \leqslant c \|\boldsymbol{v} + \boldsymbol{r}\|_{0,\Omega}$$

对任意的 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ 及任意的 $\boldsymbol{r} \in \mathbf{Ker}\ \boldsymbol{\nabla}_s$ 成立, 故

$$\|\boldsymbol{\nabla}_s \dot{\boldsymbol{v}}\|_{0,\Omega} \leqslant c \inf_{\boldsymbol{r} \in \mathbf{Ker}\ \boldsymbol{\nabla}_s} \|\boldsymbol{v} + \boldsymbol{r}\|_{1,\Omega} = c \|\dot{\boldsymbol{v}}\|_{1,\Omega}.$$

所以, 根据 Banach 开映射定理的推论 (定理 5.6-2) 即得要证的结论. □

值得注意的是, 6.15 节中的 Korn 不等式现在是定理 6.18-4 的简单推论:

定理 6.18-5 映射 $\Xi: \mathbb{E}(\Omega) \to \dot{\boldsymbol{H}}^1(\Omega)$ 是同构蕴涵着在空间 $\boldsymbol{H}^1(\Omega)$ 及 $\dot{\boldsymbol{H}}^1(\Omega)$ 中的 Korn 不等式 (定理 6.15-1 和定理 6.15-3).

证明 因为 Ξ 是同构, 故存在常数 $\dot{C}$ 使得

$$\|\Xi(\boldsymbol{e})\|_{1,\Omega} \leqslant \dot{C}\|\boldsymbol{e}\|_{0,\Omega}, \quad \text{对所有 } \boldsymbol{e} \in \mathbb{E}(\Omega),$$

或等价地使得

$$\|\dot{\boldsymbol{v}}\|_{1,\Omega} \leqslant \dot{C}\|\boldsymbol{\nabla}_s\dot{\boldsymbol{v}}\|_{0,\Omega}, \quad \text{对所有 } \dot{\boldsymbol{v}} \in \dot{\boldsymbol{H}}^1(\Omega).$$

但这正是在商空间 $\dot{\boldsymbol{H}}^1(\Omega)$ 中的 Korn 不等式, 它本身又等价于空间 $\boldsymbol{H}^1(\Omega)$ 中的 Korn 不等式 (定理 6.15-3). □

习题

6.18-1 直接证明, 对 $N = 3$ 的情况, 81 个 Saint-Venant 相容性关系式化为实际上只有 6 个独立的关系式 (不是唯一确定的).

6.18-2 当 $N = 3$ 而 Ω 不是单连通的时, 下面的问题为经典的也为弱的 Saint-Venant 引理 (定理 6.18-1 和定理 6.18-3) 提供了一个反例.

设 $\Omega := \{x = (x_1, x_2, x_3) \in \mathbb{R}^3, 1 < x_1^2 + x_2^2 < 2 \text{ 及 } 0 < x_3 < 1\} \subset \mathbb{R}^3$, 又设

$$e_{11}(x) := -\frac{x_2}{x_1^2 + x_2^2}, \quad e_{12}(x) = e_{21}(x) := \frac{x_1}{2(x_2^2 + x_2^2)},$$
$$e_{ij}(x) = 0 \quad \text{若 } i + j \geqslant 4, x \in \Omega,$$

它们满足 $\partial_{\ell i}e_{ik} + \partial_{ki}e_{j\ell} - \partial_{\ell i}e_{jk} - \partial_{kj}e_{i\ell} = 0$ 在 Ω 中. 证明, 不存在任何向量场 $\boldsymbol{v} = (v_i) \in \mathcal{C}^3(\Omega)$ 满足 $\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij}$ 在 Ω 中.

提示: 设 $\tilde{\Omega} := \Omega - \gamma$, 其中 $\gamma := \{(x_1, 0, x_3) \in \mathbb{R}^3; -2 < x_1 < -1 \text{ 及 } 0 < x_3 < 1\}$. 然后计算方程 $e_{ij}(v) = e_{ij}$ 在 $\tilde{\Omega}$ 中的显式通解 $\boldsymbol{v}$, 并证明 $\lim_{x_2 \to 0^+} +\boldsymbol{v}(x_1, x_2, x_3) \neq \lim_{x_2 \to 0^-} \boldsymbol{v}(x_1, x_2, x_3)$ 对所有 $-2 < x_1 < -1$ 及 $0 < x_3 < 1$.

6.18-3 设 Ω 是 $\mathbb{R}^3$ 中边界为 Γ 的区域, 而空间 $\mathbb{E}(\Omega)$ 及映射 $\Xi: \mathbb{E}(\Omega) \to \dot{\boldsymbol{H}}^1(\Omega)$ 如定理 6.18-4 中所定义. 给定常数 $\lambda \geqslant 0$ 和 $\mu > 0$ 以及向量场 $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$ 和 $\boldsymbol{g} \in \boldsymbol{L}^2(\Gamma)$, 定义泛函

$$j: \boldsymbol{e} \in \mathbb{E}(\Omega) \to j(\boldsymbol{e}) := \frac{1}{2}\int_\Omega \{\lambda(\mathrm{tr}\, \boldsymbol{e})^2 + 2\mu\boldsymbol{e} : \boldsymbol{e}\}dx - \ell(\Xi(\boldsymbol{e})),$$

其中泛函 $\ell: \boldsymbol{H}^1(\Omega) \to \mathbb{R}$ 定义为 $\ell(\boldsymbol{v}) := \int_\Omega \boldsymbol{f} \cdot \boldsymbol{v}dx + \int_\Gamma \boldsymbol{g} \cdot \boldsymbol{v}d\Gamma$ 对每个 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$, 并假定其满足 $\ell(\boldsymbol{r}) = 0$ 对所有 $\boldsymbol{r} \in \mathbf{Ker}\, \boldsymbol{\nabla}_s$.

(1) 证明二次极小化问题: 求 $\boldsymbol{\varepsilon} \in \mathbb{E}(\Omega)$ 使得 $j(\varepsilon) = \inf_{\boldsymbol{e} \in \mathbb{E}(\Omega)} j(\boldsymbol{e})$ 有且只有一个解 $\boldsymbol{\varepsilon}$.

(2) 证明 $\varepsilon = \boldsymbol{\nabla}_s\dot{\boldsymbol{u}}$, 其中 $\boldsymbol{u} \in \boldsymbol{H}^1(\Omega)$ 是三维线性化弹性的纯牵引问题 (定理 6.16-2) 的任一解.

注 尽管习题 6.12-2 中在空间 $\dot{\boldsymbol{H}}^1(\Omega)$ 上的极小化问题是一个具有三个未知函数的无约束问题, 而 (1) 中在空间 $\mathbb{E}(\Omega)$ 上的问题实际上却是在 $\mathbb{L}_s^2(\Omega)$ 上的具有 6 个未知函数约束二次极小

化问题, 约束就是矩阵场 $\boldsymbol{e}=(e_{ij})\in\mathbb{E}(\Omega)$ 满足的相容性关系 $\partial_{\ell j}e_{ik}+\partial_{ki}e_{j\ell}-\partial_{\ell i}e_{jk}-\partial_{kj}e_{i\ell}=0$ 在 $H^{-2}(\Omega)$ 中 (这些相容性关系实际上化为 6 个独立的; 见习题 6.18-1).

6.18-4 设 $\varepsilon_{ijk}:=1$ 或 $\varepsilon_{ijk}:=-1$ 当 $\{i,j,k\}$ 分别为 $\{1,2,3\}$ 的偶置换或奇置换时, 而 $\varepsilon_{ijk}:=0$, 如果至少有两个下标相等. 矩阵旋度算子 $\mathbf{CURL}:\mathcal{D}'(\Omega;\mathbb{M}^3)\to\mathcal{D}'(\Omega;\mathbb{M}^3)$ 和矩阵旋度旋度算子 $\mathbf{CURL\ CURL}:\mathcal{D}'(\Omega;\mathbb{M}^3)\to\mathcal{D}'(\Omega;\mathbb{S}^3)$ 分别由以下两式定义:

$$(\mathbf{CURL}\ \boldsymbol{e})_{ij}:=\varepsilon_{i\ell k}\partial_\ell e_{jk}\quad 对任意矩阵场\ \boldsymbol{e}=(e_{ij})\in\mathcal{D}'(\Omega;\mathbb{M}^3),$$

$$(\mathbf{CURL\,CURL}\ \boldsymbol{e})_{ij}:=\varepsilon_{ik\ell}\varepsilon_{jmn}\partial_{\ell n}e_{km}\quad 对任意矩阵场\ \boldsymbol{e}=(e_{ij})\in\mathcal{D}'(\Omega;\mathbb{M}^3).$$

(1) 证明, 当 $N=3$ 时 Saint-Venant 相容性关系

$$\partial_{\ell j}e_{ik}+\partial_{ki}e_{j\ell}-\partial_{\ell i}e_{jk}-\partial_{kj}e_{i\ell}=0\quad 在\ \mathcal{C}(\Omega)\ 中,\quad 对\ 1\leqslant i,j,k,\ell\leqslant 3,$$

等价于矩阵方程

$$\mathbf{CURL\,CURL}\ \boldsymbol{e}=\mathbf{0}\quad 在\ \mathcal{C}(\Omega;\mathbb{S}^3)\ 中.$$

这个式子也说明, 在这种情况下, Saint-Venant 相容性实际上可约化为只有 6 个独立的关系式.

(2) 证明, 如果 $N=3$, 则定理 6.18-2 中的 Cesàro-Volterra 路径积分公式等价于向量方程

$$\boldsymbol{v}(x)=\boldsymbol{v}(\boldsymbol{x}^0)+\boldsymbol{d}^0\wedge(\boldsymbol{x}-\boldsymbol{x}^0)+\int_{\gamma(x)}\boldsymbol{\nabla}_s\boldsymbol{v}(y)dy+\int_{\gamma(x)}(\boldsymbol{x}-\boldsymbol{y})\wedge(\mathbf{CURL}\ \boldsymbol{\nabla}_s\boldsymbol{v}(y)dy),$$

其中向量 $\boldsymbol{d}^0\in\mathbb{R}^3$ 定义为

$$\boldsymbol{d}^0:=\frac{1}{2}\begin{pmatrix}\partial_2v_3-\partial_3v_2\\ \partial_3v_1-\partial_1v_3\\ \partial_1v_2-\partial_2v_1\end{pmatrix}(\boldsymbol{x}^0).$$

6.18-5 设 Ω 是 $\mathbb{R}^N$ 的单连通开子集, $(e_{ij})\in\mathcal{C}^2(\Omega;\mathbb{S}^N)$ 是在 Ω 中满足 Saint-Venant 相容性关系的张量场.

(1) 证明, 给定任意点 $x\in\Omega$, 每一个路径积分 $\int_{\gamma_x}(e_{ij}(y))+(\partial_ke_{ij}(y)-\partial_ie_{kj}(y))(x_k-y_k)dy_j, 1\leqslant i\leqslant N$, 均与所选取的联结 x^0 和 x 的路径 $\gamma_x\in\mathcal{C}^\infty([0,1];\mathbb{R}^N)$ 无关.

提示: 像定理 6.18-1 的证明中那样, 将 Saint-Venant 关系改写为

$$\partial_\ell h_{ijk}=\partial_kh_{ij\ell}\quad 在\ \mathcal{C}(\Omega)\ 中,$$
$$其中\ h_{ijk}:=\partial_je_{ik}-\partial_ie_{jk}\in\mathcal{C}^1(\Omega).$$

然后如定理 6.17-3(b) 中那样予以证明.

(2) 利用类似于前述证明中所用的推导, 证明对每个 $x=(x_k)\in\Omega$, 由

$$v_i(x)=\int_{\gamma_x}(e_{ij}(y)+(\partial_ke_{ij}(y)-\partial_ie_{kj}(y))(x_k-y_k)dy_j,\quad 1\leqslant i\leqslant N,$$

定义的向量场 $(v_i):\Omega\to\mathbb{R}^N$ 在 Ω 中是可微的, 而其偏导数由

$$\partial_jv_i(x)=e_{ij}(x)+\int_{\gamma_x}\{\partial_je_{ik}(y)-\partial_ie_{kj}(y)\}dy_k,$$

$$\partial_iv_j=e_{ji}+\int_{\gamma_x}\{\partial_ie_{jk}(y)-\partial_je_{ki}(y)\}dy_k$$

给出. 这说明了 $\frac{1}{2}(\partial_j v_i + \partial_i v_j) = e_{ij}$ 在 Ω 中, $1 \leqslant i, j \leqslant N$.

(3) 说明如此定义的场 (v_i) 在 Ω 中是 $\mathcal{C}^3$ 类的.

6.19 J. L. Lions 引理的另一个应用: Donati 引理

整个本节中, 仍使用关于重复下标的求和约定. 我们回忆一下, 给定 $\mathbb{R}^N$ 的一个开子集 Ω, 向量散度算子 $\mathrm{div} : \mathcal{D}'(\Omega; \mathbb{M}^N) \to \mathcal{D}'(\Omega; \mathbb{R}^N)$ 定义为

$$(\mathbf{div}\, \boldsymbol{e})_i := \partial_j e_{ij}, \quad \text{对任意 } \boldsymbol{e} = (e_{ij}) \in \mathcal{D}'(\Omega; \mathbb{M}^N).$$

Saint-Venant 相容性关系 (6.18 节) 对于作为对称化梯度场的矩阵场提供了一种表征; 当然也可给出其他表征[92]. 尤其值得指出的是, Luigi Donati[93] 在 1890 年就已经注意到, 如果定义在一个开子集 $\Omega \subset \mathbb{R}^3$ 上的足够光滑的对称矩阵场 $e = (e_{ij})$ 满足

$$\int_\Omega \boldsymbol{e} : \boldsymbol{s} dx = 0, \quad \text{对所有在 } \Omega \text{ 中满足 } \mathbf{div}\, \boldsymbol{s} = \mathbf{0} \text{ 的 } \boldsymbol{s} \in \mathcal{D}(\Omega; \mathbb{S}^3),$$

则分量 e_{ij} 必须在 Ω 中满足 Saint-Venant 相容性关系. 这一结论, 再结合经典 Saint-Venant 引理 (定理 6.18-1), 就意味着如果这些关系式满足, 则存在一个光滑的向量场 $\boldsymbol{v}$ 使得 $\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e}$ 在 Ω 中 (至少当 Ω 是单连通时如此).

本节[94] 的目标是给出一些推广, 将这个经典结果拓展到具有较少正则性的矩阵场 $\boldsymbol{e}$, 仍将其称为 Donati 引理. 第一个结果是拓展到分量只属于 $H^{-1}(\Omega)$ 的对称矩阵场 $\boldsymbol{e} = (e_{ij})$, 由其得到的向量场 $\boldsymbol{v}$ (即满足 $\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e}$ 在 $\mathbb{H}^{-1}(\Omega)$ 中) 如所期望的, 属于 $L^2(\Omega)$ (定理 6.19-4); 第二和第三个拓展结果都是假设分量 e_{ij} 属于 $L^2(\Omega)$, 但不同的是, 所得到的向量场 $\boldsymbol{v}$ (即在 $\mathbb{L}^2(\Omega)$ 中满足 $\boldsymbol{e} = \boldsymbol{\nabla}_s \boldsymbol{v}$ 的向量场) 分别属于 $\boldsymbol{H}_0^1(\Omega)$ (定理 6.19-6) 及 $\boldsymbol{H}^{-1}(\Omega)$ (定理 6.19-6). 有趣的是, 这些结果成立并不要求区域是单连通的.

在下面定理中确立的算子 $\boldsymbol{\nabla}_s$ 的性质将 $H^m(\Omega)$ 中 J. L. Lions 引理 (即对梯度 $\mathbf{grad}\, v$ 属于 $\boldsymbol{H}^m(\Omega)$ 的 $H^m(\Omega)$ 中分布 v 而言的; 参见定理 6.11-5) 拓展到对称化梯度属于 $\mathbb{H}^m(\Omega)$ 的向量场 $\boldsymbol{v} \in \boldsymbol{H}^m(\Omega)$. 像本节中的其他结果一样, 这进一步说明, 矩

---

[92] 关于作为对称化梯度场的矩阵场经典表征这一问题的产生和发展历史, 可参见:

M. E. GURTIN [1972]: The linear theory of elasticity. Handbuch der Physik, Volume VIa/2 (S. FLÜGGE & C. TRUESDELL, editors), pp. 1–295, Springer, Berlin.

[93] L. DONATI [1890]: Illustrazione al teorema del Menabrea. Memorie della Accademia delle Scienze dell'Istituto di Bologna **10**, 267–274.

L. DONATI [1894]: Ulteriori osservazioni intorno al teorema del Menabrea. Memorie della Accademia delle Scienze dell'Istituto di Bologna **4**, 449–474.

[94] 本节的内容取自:

C. AMROUCHE; P. G. CIARLET; L. GRATIE; S. KESAVAN [2006]: On the characterization of matrix fields as linearized strain tensor fields. Journal de Mathématiques Pures et Appliquées **86**, 116–132.

阵算子 $\nabla_s : \mathcal{D}'(\Omega;\mathbb{R}^N) \to \mathcal{D}'(\Omega;\mathbb{S}^N)$ 实际上是向量算子 $\mathbf{grad} : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega;\mathbb{R}^N)$ 的"矩阵类似".

定理 6.19-1 ($\boldsymbol{H}^m(\Omega)$ 中的 J. L. Lions 引理: 向量形式) 设 Ω 是 $\mathbb{R}^N$ 中的区域, $m \in \mathbb{Z}$. 则

$$\boldsymbol{v} \in \boldsymbol{H}^m(\Omega) \text{ 和 } \nabla_s \boldsymbol{v} \in \mathbb{H}^m(\Omega) \text{ 意味着 } \boldsymbol{v} \in \boldsymbol{H}^{m+1}(\Omega).$$

证明 我们记得, $H^m(\Omega)$ 中的 J. L. Lions 引理 (定理 6.11-5) 断言, 对任何 $m \in \mathbb{Z}, v \in H^m(\Omega)$ 及 $\mathbf{grad}\, v \in H^m(\Omega)$ 意味着 $v \in H^{m+1}(\Omega)$.

设 $\boldsymbol{v} = (v_i) \in \boldsymbol{H}^m(\Omega)$ 使得 $\nabla_s \boldsymbol{v} \in \mathbb{H}^m(\Omega)$ 对某整数 $m \in \mathbb{Z}$. 则恒等式

$$(\mathbf{grad}(\partial_k v_i))_j = \partial_j((\nabla_s \boldsymbol{v})_{ik}) + \partial_k((\nabla_s \boldsymbol{v})_{ij}) - \partial_i((\nabla_s \boldsymbol{v})_{jk})$$

说明, 梯度矩阵 $\nabla \boldsymbol{v}$ 的每个分量 $\partial_k v_i \in H^{m-1}(\Omega)$ 也满足 $\mathbf{grad}(\partial_k v_i) \in \boldsymbol{H}^{m-1}(\Omega)$, 所以, $H^{m-1}(\Omega)$ 中的 J. L. Lions 引理说明 $\partial_k v_i \in H^m(\Omega)$. 由假设 $v_i \in H^m(\Omega)$, 这次在空间 $H^m(\Omega)$ 中再用同一个 J. L. Lions 引理, 说明 $v_i \in H^{m+1}(\Omega)$. □

注 与原本的 J. L. Lions 引理相比较, 上述 J. L. Lions 引理的向量形式在 $m = 0$ 的情况, 不再是平凡的了.

在下面定理中, 将算子 ∇_s 视为从空间 $\boldsymbol{L}^2(\Omega)$ 作用到空间 $\mathbb{H}^{-1}(\Omega)$, 给出了两条性质. 注意, 这些性质正是在定理 6.14-1 的 (a) 及 (b) 中对于向量场所确立的性质之自然的矩阵类似.

定理 6.19-2 设 Ω 是 $\mathbb{R}^N$ 中的区域. 则

(a) 连续算子

$$\nabla_s : \boldsymbol{L}^2(\Omega) \to \mathbb{H}^{-1}(\Omega)$$

的对偶是连续算子

$$-\mathbf{div} : \mathbb{H}_0^1(\Omega) \to \boldsymbol{L}^2(\Omega).$$

(b) 空间 $\boldsymbol{L}^2(\Omega)$ 在算子 ∇_s 作用下的像在 $\mathbb{H}^{-1}(\Omega)$ 中是闭的.

证明 (i) 对任意的 $\boldsymbol{v} = (v_i) \in \boldsymbol{L}^2(\Omega)$ 和任意的 $\boldsymbol{e} = (e_{ij}) \in \mathbb{H}_0^1(\Omega)$,

$$\begin{aligned} {}_{\mathbb{H}^{-1}(\Omega)}\langle \nabla_s \boldsymbol{v}, \boldsymbol{e} \rangle_{\mathbb{H}_0^1(\Omega)} &= {}_{H^{-1}(\Omega)}\langle \partial_j v_i, e_{ij} \rangle_{H_0^1(\Omega)} \\ &= {}_{L^2(\Omega)}\langle v_i, -\partial_j e_{ij} \rangle_{L^2(\Omega)} = {}_{\boldsymbol{L}^2(\Omega)}\langle \boldsymbol{v}, -\mathbf{div}\, \boldsymbol{e} \rangle_{\boldsymbol{L}^2(\Omega)} \end{aligned}$$

(在第一个等式中利用了 $\boldsymbol{e}$ 的对称性). 所以 $\nabla_s : \boldsymbol{L}^2(\Omega) \to \mathbb{H}^{-1}(\Omega)$ 的对偶算子是 $-\mathbf{div} : \mathbb{H}_0^1(\Omega) \to \boldsymbol{L}^2(\Omega)$ 以及 $\nabla_s : \dot{\boldsymbol{L}}^2(\Omega) \to \mathbb{H}^{-1}(\Omega)$ (其定义为: 对每个 $\dot{\boldsymbol{v}} \in \dot{\boldsymbol{L}}^2(\Omega), \nabla_s \dot{\boldsymbol{v}} := \nabla_s \boldsymbol{w}$ 对任意 $\boldsymbol{w} \in \dot{\boldsymbol{v}}$) 的对偶是 $-\mathbf{div} : \mathbb{H}_0^1(\Omega) \to \dot{\boldsymbol{L}}^2(\Omega)$. 这就证明了 (a).

(ii) 存在常数 C 使得

$$\|\boldsymbol{v}\|_{0,\Omega} \leqslant C(\|\boldsymbol{v}\|_{-1,\Omega} + \|\boldsymbol{\nabla}_s \boldsymbol{v}\|_{-1,\Omega}), \quad \text{对所有 } \boldsymbol{v} \in \boldsymbol{L}^2(\Omega).$$

容易看出, 空间

$$\boldsymbol{K}(\Omega) := \{\boldsymbol{v} \in \boldsymbol{H}^{-1}(\Omega); \boldsymbol{\nabla}_s \boldsymbol{v} \in \mathbb{H}^{-1}(\Omega)\},$$

装备以下范数

$$\boldsymbol{v} \in \boldsymbol{K}(\Omega) \to \|\boldsymbol{v}\|_{\boldsymbol{K}(\Omega)} := (\|\boldsymbol{v}\|_{-1,\Omega}^2 + \|\boldsymbol{\nabla}_s \boldsymbol{v}\|_{-1,\Omega}^2)^{1/2},$$

是完备的 (类似于定理 6.14-1 的证明 (iii)). 恒等映射 $\iota : (\boldsymbol{L}^2(\Omega), \|\cdot\|_{0,\Omega}) \to (\boldsymbol{K}(\Omega), \|\cdot\|_{\boldsymbol{K}(\Omega)})$ 是单射、连续的 (相应的不等式显然成立), 根据 $\boldsymbol{H}^{-1}(\Omega)$ 中的 J. L. Lions 引理 (定理 6.19-1) 还是满射; Banach 开映射定理的推论说明, ι 的逆映射也是连续的. 这正是 (ii) 中所示要证明的不等式.

(iii) 存在常数 $\dot{C}$ 使得

$$\|\dot{\boldsymbol{v}}\|_{0,\Omega} \leqslant \dot{C}\|\boldsymbol{\nabla}_s \dot{\boldsymbol{v}}\|_{-1,\Omega}, \quad \text{对所有 } \dot{\boldsymbol{v}} \in \dot{\boldsymbol{L}}^2(\Omega).$$

假若这个不等式不成立, 则存在 $\dot{\boldsymbol{v}}^k \in \dot{\boldsymbol{L}}^2(\Omega), k \geqslant 1$, 使得

$$\|\dot{\boldsymbol{v}}^k\|_{0,\Omega} = 1 \quad \text{对所有 } k \geqslant 1 \text{ 及 } \|\boldsymbol{\nabla}_s \dot{\boldsymbol{v}}^k\|_{-1,\Omega} \to 0 \quad \text{当 } k \to \infty \text{ 时}.$$

因空间 $\mathbf{Ker}\ \boldsymbol{\nabla}_s$ 是有限维的 (定理 6.15-2), 故对每个 $\dot{\boldsymbol{v}} \in \dot{\boldsymbol{L}}^2(\Omega)$, 存在一个元素 $\boldsymbol{w} \in \dot{\boldsymbol{v}}$ 使得 $\|\boldsymbol{w}\|_{0,\Omega} = \|\dot{\boldsymbol{v}}\|_{0,\Omega} := \inf_{\boldsymbol{r} \in \mathbf{Ker}\ \boldsymbol{\nabla}_s} \|\boldsymbol{v} + \boldsymbol{r}\|_{0,\Omega}$. 所以对每个整数 $k \geqslant 1$, 存在 $\boldsymbol{w}^k \in \dot{\boldsymbol{v}} \subset \boldsymbol{L}^2(\Omega)$ 使得

$$\|\boldsymbol{w}^k\|_{0,\Omega} = 1 \quad \text{对所有 } k \geqslant 1 \text{ 及 } \|\boldsymbol{\nabla}_s \boldsymbol{w}^k\|_{-1,\Omega} = \|\boldsymbol{\nabla}_s \dot{\boldsymbol{v}}^k\|_{-1,\Omega} \to 0 \quad \text{当 } k \to \infty \text{ 时}.$$

根据空间 $L^2(\Omega)$ 中的 Rellich-Kondrachov 紧嵌入定理 (定理 6.11-3), 存在一个在 $\boldsymbol{H}^{-1}(\Omega)$ 中收敛的子列 $(\boldsymbol{w}^{\sigma(k)})_{k=1}^\infty$. 由于子列 $(\boldsymbol{\nabla}_s \boldsymbol{w}^{\sigma(k)})_{k=1}^\infty$ 在 $\mathbb{H}^{-1}(\Omega)$ 中收敛 (于 0, 但这个结果在现阶段不需要), 子列 $(\boldsymbol{w}^{\sigma(k)})_{k=1}^\infty$ 因此在空间 $(\boldsymbol{K}(\Omega), \|\cdot\|_{\boldsymbol{K}(\Omega)})$ 中是 Cauchy 序列. 因此由 (ii) 中所确立的不等式得这个子列也是空间 $\boldsymbol{L}^2(\Omega)$ 中的 Cauchy 序列.

所以, 存在 $\boldsymbol{w} \in \boldsymbol{L}^2(\Omega)$ 使得

$$\boldsymbol{w}^{\sigma(k)} \xrightarrow[k \to \infty]{} \boldsymbol{w} \quad \text{在 } \boldsymbol{L}^2(\Omega) \text{ 中}.$$

此外

$$\boldsymbol{\nabla}_s \boldsymbol{w}^{(k)} \xrightarrow[k \to \infty]{} \boldsymbol{0} = \boldsymbol{\nabla}_s \boldsymbol{w} \quad \text{在 } \boldsymbol{H}^{-1}(\Omega) \text{ 中},$$

这意味着 $\boldsymbol{w} \in \mathbf{Ker}\ \boldsymbol{\nabla}_s$. 因此

$$\dot{\boldsymbol{w}}^{\sigma(k)} \xrightarrow[k\to\infty]{} \dot{\boldsymbol{w}} = \dot{\mathbf{0}} \quad \text{在 } \dot{\boldsymbol{L}}^2(\Omega) \text{ 中},$$

这与 $\|\dot{\boldsymbol{w}}^{\sigma(k)}\|_{0,\Omega} = \|\boldsymbol{w}^{\sigma(k)}\|_{0,\Omega} = 1$ 对所有 $k \geqslant 1$ 矛盾. 故所宣示的不等式成立.

(iv) 显然, 在算子 $\boldsymbol{\nabla}_s$ 的作用下, 空间 $\boldsymbol{L}^2(\Omega)$ 与 $\dot{\boldsymbol{L}}^2(\Omega)$ 在空间 $\mathbb{H}^{-1}(\Omega)$ 中的像是等同的. 线性算子 $\boldsymbol{\nabla}_s : \dot{\boldsymbol{L}}^2(\Omega) \to \mathbb{H}^{-1}(\Omega)$ 是单射 (因 $\dot{\boldsymbol{L}}^2(\Omega) = \boldsymbol{L}^2(\Omega)/\mathbf{Ker}\ \boldsymbol{\nabla}_s$), 显然还是连续的, 而且有从 $\mathbf{Im}\ \boldsymbol{\nabla}_s \subset \mathbb{H}^{-1}(\Omega)$ 到 $\dot{\boldsymbol{L}}^2(\Omega)$ 上的逆, 由 (iii) 知这个逆也是连续的. 因此空间 $\mathbf{Im}\ \boldsymbol{\nabla}_s$ 是 $\mathbb{H}^{-1}(\Omega)$ 的一个完备子空间, 当然也是 $\mathbb{H}^{-1}(\Omega)$ 的一个闭子空间. (b) 得证. □

顺便提一下, (ii) 中确立的不等式构成在 $\boldsymbol{L}^2(\Omega)$ 中的 Korn 不等式.

在下面定理中, 将算子 $\boldsymbol{\nabla}_s$ 视为从空间 $\boldsymbol{H}_0^1(\Omega)$ 作用到空间 $\mathbb{L}^2(\Omega)$, 给出了两条性质; 要注意, 在这里因为 $\mathbf{Ker}\ \boldsymbol{\nabla}_s = \mathbf{0}$, $\boldsymbol{\nabla}_s$ 现在变成单射.

定理 6.19-3 设 Ω 是 $\mathbb{R}^N$ 中的区域. 则:

(a) 单射的连续算子

$$\boldsymbol{\nabla}_s : \boldsymbol{H}_0^1(\Omega) \to \mathbb{L}^2(\Omega)$$

的对偶是连续算子

$$-\mathbf{div} : \mathbb{L}^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega).$$

(b) 空间 $\boldsymbol{H}_0^1(\Omega)$ 在算子 $\boldsymbol{\nabla}_s$ 作用下的像在 $\mathbb{L}^2(\Omega)$ 中是闭的.

证明 证明与定理 6.19-2 中的类似, 甚至于更简单些, 这是因为初等的计算给出 (习题 6.15-1)

$$|\boldsymbol{v}|_{1,\Omega} := \left(\sum_{i,j} \|\partial_j v_i\|_{0,\Omega}^2\right) \leqslant \sqrt{2}\|\boldsymbol{\nabla}_s \boldsymbol{v}\|_{0,\Omega}, \quad \text{对所有 } \boldsymbol{v} = (v_i) \in \boldsymbol{H}_0^1(\Omega).$$

这个关系式意味着存在常数 C 使得 (定理 6.5-2)

$$\|\boldsymbol{v}\|_{1,\Omega} \leqslant C\|\boldsymbol{\nabla}_s \boldsymbol{v}\|, \quad \text{对所有 } \boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega),$$

这个不等式就与定理 6.19-2 的证明中 (iii) 类似. □

现在我们可以来证明第一个 Donati 引理, 这一引理是定理 6.14-2 的 “矩阵类似”.

定理 6.19-4 ($L^2(\Omega)$ 中的 Donati 引理) 设 Ω 是 $\mathbb{R}^N$ 中的区域. 给定一矩阵场 $\boldsymbol{e} \in \mathbb{H}^{-1}(\Omega)$, 存在一个向量场 $\boldsymbol{v}$ 使得

$$\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) \quad \text{及} \quad \boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e} \text{ 在 } \mathbb{H}^{-1}(\Omega) \text{ 中}$$

的充分必要条件是

$$_{\mathbb{H}^{-1}(\Omega)}\langle \boldsymbol{e}, \boldsymbol{s}\rangle_{\mathbb{H}_0^1(\Omega)} = 0 \quad \text{对所有满足 } \mathbf{div}\, \boldsymbol{s} = \mathbf{0} \text{ 在 } \boldsymbol{L}^2(\Omega) \text{ 中的 } \boldsymbol{s} \in \mathbb{H}_0^1(\Omega).$$

方程 $\nabla_s \tilde{\boldsymbol{v}} = \boldsymbol{e}$ 的所有其他解 $\tilde{\boldsymbol{v}} \in \boldsymbol{L}^2(\Omega)$ 具有以下形式:

$$\tilde{\boldsymbol{v}}(x) = \boldsymbol{v}(x) + \boldsymbol{B}x + \boldsymbol{c}, \quad \text{对几乎所有 } x \in \Omega,$$

其中 $\boldsymbol{B}$ 是某 $N \times N$ 反对称矩阵, $\boldsymbol{c} \in \mathbb{R}^N$ 为某向量.

证明　在定理 6.19-2 中已经证明, $\nabla_s : \boldsymbol{L}^2(\Omega) \to \mathbb{H}^{-1}(\Omega)$ 的对偶是 $-\mathbf{div} : \mathbb{H}_0^{-1}(\Omega) \to \boldsymbol{L}^2(\Omega), \boldsymbol{L}^2(\Omega)$ 在 ∇_s 作用下的像 $\mathrm{Im}\,\nabla_s$ 在 $\mathbb{H}^{-1}(\Omega)$ 中是闭的. 因此, Banach 闭值域定理 (第一部分; 见定理 5.11-5) 意味着有

$$\mathrm{Im}\,\nabla_s = \{\boldsymbol{e} \in \mathbb{H}^{-1}(\Omega);\, {}_{\mathbb{H}^{-1}(\Omega)}\langle \boldsymbol{e}, \boldsymbol{s}\rangle_{\mathbb{H}_0^1(\Omega)} = 0 \text{ 对所有 } \boldsymbol{s} \in \mathrm{Ker}(-\mathbf{div})\},$$

这正是定理中所要证明的. 而方程 $\nabla_s \tilde{\boldsymbol{v}} = \boldsymbol{e}$ 的所有其他解 $\tilde{\boldsymbol{v}} \in \boldsymbol{L}^2(\Omega)$ 具有定理中指定的形式结论, 可由定理 6.15-2 中确立的关于空间 $\mathbf{Ker}\,\nabla_s$ 的表征得到. □

注　定理 6.19-4 可以推广[95] 到矩阵场 $\boldsymbol{e} \in \mathbb{W}^{-1,p}(\Omega), 1 < p < \infty$, 它满足

$${}_{\mathbb{W}^{-1,p}(\Omega)}\langle \boldsymbol{e}, \boldsymbol{s}\rangle_{\mathbb{W}_0^{1,q}(\Omega)} = 0$$

对所有满足 $\mathbf{div}\,\boldsymbol{s} = \mathbf{0}$ 在 $\boldsymbol{L}^q(\Omega)$ 中的 $\boldsymbol{s} \in \mathbb{W}_0^{1,q}(\Omega)$, 其中 q 是 p 的共轭指数.

上面的 Donati 引理其实是定理 6.19-2 的一个推论. 其他的 Donati 引理可类似地得到, 这次是作为定理 6.19-3 的推论.

定理 6.19-5 ($\boldsymbol{H}_0^1(\Omega)$ 中的 Donati 引理)　*设 Ω 是 $\mathbb{R}^N$ 中的区域. 给定一个矩阵场 $\boldsymbol{e} \in \mathbb{L}^2(\Omega)$, 存在一个向量场 $\boldsymbol{v}$ 使得*

$$\boldsymbol{v} \in \boldsymbol{H}_0^1(\Omega) \quad \text{及} \quad \nabla_s \boldsymbol{v} = \boldsymbol{e} \text{ 在 } \mathbb{L}^2(\Omega) \text{ 中}$$

的充分必要条件是

$$\int_\Omega \boldsymbol{e} : \boldsymbol{s} dx = 0 \quad \text{对所有在 } \boldsymbol{H}^{-1}(\Omega) \text{ 中满足 } \mathbf{div}\,\boldsymbol{s} = \mathbf{0} \text{ 的 } \boldsymbol{s} \in \mathbb{L}^2(\Omega).$$

在这种情况下, 向量场 $\boldsymbol{v}$ 是唯一的.

证明　由于 $\nabla_s : \boldsymbol{H}_0^1(\Omega) \to \mathbb{L}^2(\Omega)$ 的对偶算子是 $-\mathbf{div} : \mathbb{L}^2(\Omega) \to \boldsymbol{H}^{-1}(\Omega)$, 而且 $\boldsymbol{H}_0^1(\Omega)$ 在 ∇_s 作用下的像在 $\mathbb{L}^2(\Omega)$ 中是闭的 (定理 6.19-3), 故向量场 $\boldsymbol{v}$ 的存在性可由 Banach 闭值域定理得到 (类似于定理 6.19-4 的证明, 但在此要讨论的算子 ∇_s 是从 $\boldsymbol{H}_0^1(\Omega)$ 作用到 $\mathbb{L}^2(\Omega)$ 中). 在这种情况下, $\mathbf{Ker}\,\nabla_s = \{\mathbf{0}\}$, 这意味着这种向量场 $\boldsymbol{v}$ 是唯一的. □

---

[95] G. Geymonat; F. Krasucki [2005]: Some remarks on the compatibility conditions in elasticity. Accademia Nazionale delle Scienze detta dei XL. Rendiconti. Serie V. Memorie di Matematica e Applicazioni. Parte I, **29**, 175–181.

注 类似的结果对更一般的边界条件成立[96), 即 $\boldsymbol{v} = \mathbf{0}$ 在 $\Gamma := \partial\Omega$ 的一个相对开子集 Γ_0 上, 且 $d\Gamma\text{-meas-}\Gamma_0 > 0$.

最后, 第三个 Donati 引理可以作为 J. L. Lions 引理的向量形式 (定理 6.19-1) 以及第一 Donati 引理 (定理 6.19-4) 的一个推论而导出. 要注意, 满足 $\mathbf{div}\,\boldsymbol{s} = \mathbf{0}$ 的张量场 $\boldsymbol{s}$ 现在的范围是在空间 $\mathbb{H}_0^1(\Omega)$ 中, 而不是像在定理 6.19-5 中那样在空间 $\mathbb{L}^2(\Omega)$ 中; 而作为结果, 所找到的向量场 $\boldsymbol{v}$ 现在在空间 $\boldsymbol{H}^1(\Omega)$ 中, 而不是空间 $\boldsymbol{H}_0^1(\Omega)$.

定理 6.19-6 ($\boldsymbol{H}^1(\Omega)$ 中的 Donati 引理) *设 Ω 是 $\mathbb{R}^N$ 中的一区域. 给定矩阵场 $\boldsymbol{e} \in \mathbb{L}^2(\Omega)$, 存在向量场 $\boldsymbol{v}$ 使得*

$$\boldsymbol{v} \in \boldsymbol{H}^1(\Omega) \quad \text{及} \quad \boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e} \text{ 在 } \mathbb{L}^2(\Omega) \text{ 中}$$

的充分必要条件是

$$\int_\Omega \boldsymbol{e} : \boldsymbol{s} dx = 0 \quad \text{对所有在 } \boldsymbol{L}^2(\Omega) \text{ 中满足 } \mathbf{div}\,\boldsymbol{s} = \mathbf{0} \text{ 的 } \boldsymbol{s} \in \mathbb{H}_0^1(\Omega).$$

方程 $\boldsymbol{\nabla}_s \tilde{\boldsymbol{v}} = \boldsymbol{e}$ 的所有其他解 $\tilde{\boldsymbol{v}} \in \boldsymbol{L}^2(\Omega)$ 具有以下形式:

$$\tilde{\boldsymbol{v}}(x) = \boldsymbol{v}(x) + \boldsymbol{B}x + \boldsymbol{c}, \quad \text{对几乎所有 } x \in \Omega,$$

其中 $\boldsymbol{B}$ 是某 $N \times N$ 反对称矩阵, $\boldsymbol{c} \in \mathbb{R}^N$ 是某一向量.

证明 设 $\boldsymbol{e} \in \mathbb{L}^2(\Omega)$ 使得 $\int_\Omega \boldsymbol{e} : \boldsymbol{s} dx = 0$ 对所有在 $\boldsymbol{L}^2(\Omega)$ 中满足 $\mathbf{div}\,\boldsymbol{s} = \mathbf{0}$ 的 $\boldsymbol{s} \in \mathbb{H}_0^1(\Omega)$ 成立.

由于 $\mathbb{L}^2(\Omega) \subset \mathbb{H}^{-1}(\Omega)$, 定理 6.19-4 说明, 存在 $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega)$ 使得 $\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e}$ 在 $\mathbb{H}^{-1}(\Omega)$ 中; 由于假设 $\boldsymbol{e} \in \mathbb{L}^2(\Omega)$, 故 $\boldsymbol{\nabla}_s \boldsymbol{v} \in \mathbb{L}^2(\Omega)$. 由定理 6.19-1 在 $m = 0$ 的情况得, $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$. 所以定理中显示的关系是充分的.

反之, 对某 $\boldsymbol{v} = (v_i) \in \boldsymbol{H}^1(\Omega)$, 假设 $\boldsymbol{e} = (e_{ij}) = \boldsymbol{\nabla}_s \boldsymbol{v}$. 则由 $\boldsymbol{e}$ 的对称性及 Green 公式 (定理 6.6-7) 有

$$\begin{aligned}\int_\Omega \boldsymbol{e} : \boldsymbol{s} dx &= \int_\Omega e_{ij} s_{ij} dx = \int_\Omega (\partial_j v_i) s_{ij} dx \\ &= -\int_\Omega v_i \partial_j s_{ij} dx = -\int_\Omega \boldsymbol{v} \cdot \mathbf{div}\,\boldsymbol{s} dx, \quad \text{对所有 } \boldsymbol{s} = (s_{ij}) \in \mathbb{H}_0^1(\Omega).\end{aligned}$$

所以如果 $\boldsymbol{s} \in \mathbb{H}_0^1(\Omega)$ 满足 $\mathbf{div}\,\boldsymbol{s} = \mathbf{0}$ 在 $\mathbb{L}^2(\Omega)$ 中, 则 $\int_\Omega \boldsymbol{e} : \boldsymbol{s} dx = 0$. 故所显示的关系也是必要的.

由定理 6.15-2 立得非唯一性结果. □

在习题 6.19-1 中, 给出了定理 6.19-4 和 6.19-6 的一些值得关注的补充.

[96) G. GEYMONAT; F. KRASUCKI [2005]: Some remarks on the compatibility conditions in elasticity. Accademia Nazionale delle Scienze detta dei XL. Rendiconti. Serie V. Memorie di Matematica e Applicazioni. Parte I, **29**, 175–181.

习题

6.19-1 给定 $\mathbb{R}^N$ 中的区域 Ω, 定义空间

$$\mathbb{V}(\Omega) := \{\boldsymbol{s} \in \mathbb{H}_0^1(\Omega); \mathbf{div}\, \boldsymbol{s} = \mathbf{0} \text{ 在 } \Omega \text{ 中}\} \quad \text{和}$$

$$\mathbb{W}(\Omega) := \{\boldsymbol{\sigma} \in \mathbb{D}(\Omega); \mathbf{div}\, \boldsymbol{\sigma} = \mathbf{0} \text{ 在 } \Omega \text{ 中}\}.$$

(1) 设矩阵场 $\boldsymbol{e} \in \mathbb{H}^{-1}(\Omega)$ 满足

$${}_{\mathbb{H}^{-1}(\Omega)}\langle \boldsymbol{e}, \boldsymbol{s}\rangle_{\mathbb{H}_0^1(\Omega)} = 0 \quad \text{对所有 } \boldsymbol{s} \in \mathbb{W}(\Omega).$$

证明存在向量场 $\boldsymbol{v} \in \boldsymbol{L}^2(\Omega)$ 使得 $\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e}$ 在 $\mathbb{H}^{-1}(\Omega)$ 中.

提示: 参见习题 6.14-1 中给出的提示.

(2) 利用 (1) 和定理 4.3-2 证明, $\mathbb{V}(\Omega)$ 的子空间 $\mathbb{W}(\Omega)$ 在 $(\mathbb{V}(\Omega), |\cdot|_{1,\Omega})$ 中是稠密的.

(3) 设矩阵场 $\boldsymbol{e} \in \mathbb{L}^2(\Omega)$ 满足

$$\int_\Omega \boldsymbol{e} : \boldsymbol{\sigma} dx = 0, \quad \text{对所有 } \boldsymbol{\sigma} \in \mathbb{W}(\Omega).$$

证明, 存在一个向量场 $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega)$ 使得 $\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e}$ 在 $\mathbb{L}^2(\Omega)$ 中[97].

注 实际上可以证明下述结果[98] 更一般地在分布意义下成立: 设 Ω 是 $\mathbb{R}^N$ 的任一开子集. 如果一矩阵场 $\boldsymbol{e} = (e_{ij}) \in \mathbb{D}'(\Omega)$ 使得 ${}_{\mathbb{D}'(\Omega)}\langle \boldsymbol{e}, \boldsymbol{\sigma}\rangle_{\mathbb{D}(\Omega)} := {}_{\mathcal{D}'(\Omega)}\langle e_{ij}, \sigma_{ij}\rangle_{\mathcal{D}(\Omega)} = 0$ 对所有在 Ω 中满足 $\mathbf{div}\, \boldsymbol{\sigma} = \mathbf{0}$ 的矩阵场 $\boldsymbol{\sigma} = (\sigma_{ij}) \in \mathbb{D}(\Omega)$ 成立, 则存在一个向量场 $\boldsymbol{v} \in \mathcal{D}'(\Omega)$ 满足 $\boldsymbol{\nabla}_s \boldsymbol{v} = \boldsymbol{e}$ 在 $\mathbb{D}'(\Omega)$ 中.

6.19-2 证明空间 $\mathbb{V}(\Omega) := \{\boldsymbol{s} \in \mathbb{H}_0^1(\Omega); \mathbf{div}\, \boldsymbol{s} = \mathbf{0} \text{ 在 } \Omega \text{ 中}\}$ (与习题 6.19-1 中的相同) 关于范数 $\|\cdot\|_{0,\Omega}$ 的闭包是空间 $\{\boldsymbol{s} \in \mathbb{L}^2(\Omega); \mathbf{div}\, \boldsymbol{s} = \mathbf{0} \text{ 在 } \boldsymbol{H}^{-1}(\Omega) \text{ 中}\}$ 的严格子空间 (当然, 同样的结论对空间 $\{\boldsymbol{\sigma} \in \mathbb{D}(\Omega); \mathbf{div}\, \boldsymbol{\sigma} = \mathbf{0} \text{ 在 } \Omega \text{ 中}\}$ 的闭包更应成立).

6.19-3 设 Ω 为 $\mathbb{R}^3$ 中边界为 Γ 的一区域. 定义 Hilbert 空间

$$\tilde{\mathbb{E}}(\Omega) := \left\{\boldsymbol{e} \in \mathbb{L}^2(\Omega); \int_\omega \boldsymbol{e} : \boldsymbol{s} dx = 0 \text{ 对所有在 } \boldsymbol{H}^{-1}(\Omega) \text{ 中满足 } \mathbf{div}\, \boldsymbol{s} = \mathbf{0} \text{ 的 } \boldsymbol{s} \in \mathbb{L}^2(\Omega)\right\},$$

而且对任一 $\boldsymbol{e} \in \tilde{\mathbb{E}}(\Omega)$, 令 $\tilde{\Xi}(\boldsymbol{e})$ 表示在空间 $\boldsymbol{H}_0^1(\Omega)$ 中满足 $\boldsymbol{\nabla}_s \tilde{\Xi}(\boldsymbol{e}) = \boldsymbol{e}$ 的唯一元素 (定理 6.19-5).

(1) 证明, 以上述方式定义的线性算子 $\tilde{\Xi} : \tilde{\mathbb{E}}(\Omega) \to \boldsymbol{H}_0^1(\Omega)$ 是双射、连续的, 而且有连续的逆.

(2) 给定常数 $\lambda \geqslant 0$ 和 $\mu > 0$ 以及一个向量场 $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$, 并定义泛函

$$\tilde{j} : \boldsymbol{e} \in \tilde{\mathbb{E}}(\Omega) \to \tilde{j}(\boldsymbol{e}) := \frac{1}{2} \int_\Omega \{\lambda(\operatorname{tr} \boldsymbol{e})^2 + 2\mu \boldsymbol{e} : \boldsymbol{e}\} dx - \int_\Omega \boldsymbol{f} \cdot \tilde{\Xi}(\boldsymbol{e}) dx.$$

证明下列二次极小化问题: 求 $\tilde{\varepsilon} \in \tilde{\mathbb{E}}(\Omega)$ 使得 $\tilde{j}(\tilde{\varepsilon}) = \inf_{\boldsymbol{e} \in \tilde{\mathbb{E}}(\Omega)} \tilde{j}(\boldsymbol{e})$ 有且仅有一个解.

(3) 证明 $\tilde{\boldsymbol{\varepsilon}} = \boldsymbol{\nabla}_s \boldsymbol{u}$, 其中 $\boldsymbol{u} \in \boldsymbol{H}_0^1(\Omega)$ 是线性化弹性的纯位移问题 (6.16 节) 的解.

---

[97] 这个结果属于:

T. W. TING [1974]: St. Venant's compatibility conditions. Tensors, N. S. **28**, 5–12.

[98] J. J. MOREAU [1979]: Duality characterization of strain tensor distributions in an arbitrary open set. Journal of Mathematical Analysis and Applications **72**, 760–770.

注 将此题与习题 6.18-3 相比较将会很有收益.

6.20 Pfaff 方程组

我们以讨论特定的一类一阶线性偏微分方程组来结束这一章. 这个方程组 (连同经典 Poincaré 引理; 参阅定理 6.17-2) 在关于 $\mathbb{R}^n$ 中开子集的 Riemann 几何基本定理 (定理 8.6-1) 的证明, 以及曲面论基本定理 (定理 8.16-1) 的证明中起着关键的作用.

设 Ω 是 $\mathbb{R}^N$ 的一个开子集, $n \geqslant 1$ 是整数. Pfaff 方程组[99] 是形如

$$\partial_i \boldsymbol{F}(x) = \boldsymbol{F}(x)\boldsymbol{\Gamma}_i(x), \quad x \in \Omega, 1 \leqslant i \leqslant N$$

的 N 个方程的集合, 其中矩阵场 $\boldsymbol{\Gamma}_i : \Omega \to \mathbb{M}^n, 1 \leqslant i \leqslant N$, 是给定的, 而未知量是矩阵场 $\boldsymbol{F} : \Omega \to \mathbb{M}^n$. 未知的矩阵场通常要求满足形如

$$\boldsymbol{F}(x^0) = \boldsymbol{F}^0$$

的附加条件, 其中点 $x^0 \in \Omega$, 而矩阵 $\boldsymbol{F}^0 \in \mathbb{M}^n$ 是给定的 (如果 $n = 1$, 这个条件就是一个常微分方程的初始条件), 这个条件保证解的唯一性.

注 Pfaff 方程可取更一般的形式; 见习题 6.20-1.

对这种 Pfaff 方程组解的存在性可立即给出一个必要条件, 这实际上表示解的二阶偏导数的可交换性 (正像在 6.17 节中, 对于方程组 $\partial_i p(x) = h_i(x), x \in \Omega, 1 \leqslant i \leqslant N$, 的必要条件): 假设 $\boldsymbol{\Gamma}_i \in \mathcal{C}^1(\Omega; \mathbb{M}^n), 1 \leqslant i \leqslant N$; 则解 $\boldsymbol{F} \in \mathcal{C}^2(\Omega; \mathbb{M}^n)$ 必须满足 $\partial_{ki}\boldsymbol{F} = \partial_{ik}\boldsymbol{F}, x \in \Omega, 1 \leqslant i, k \leqslant N$, 即

$$\boldsymbol{F}(x)(\boldsymbol{\Gamma}_i(x)\boldsymbol{\Gamma}_k(x) + \partial_i\boldsymbol{\Gamma}_k(x)) = \boldsymbol{F}(x)(\boldsymbol{\Gamma}_k(x)\boldsymbol{\Gamma}_i(x) + \partial_k\boldsymbol{\Gamma}_i(x)), \quad x \in \Omega, 1 \leqslant i, k \leqslant N.$$

这样, 在矩阵 $\boldsymbol{F}(x) \in \mathbb{M}^n$ 对每个 $x \in \Omega$ 都是可逆的假设下, 得

$$\partial_i\boldsymbol{\Gamma}_k(x) - \partial_k\boldsymbol{\Gamma}_i(x) + \boldsymbol{\Gamma}_i(x)\boldsymbol{\Gamma}_k(x) - \boldsymbol{\Gamma}_k(x)\boldsymbol{\Gamma}_i(x) = 0, \quad x \in \Omega, 1 \leqslant i, k \leqslant N.$$

值得关注的是, 我们下面将要证明, 如果开集 Ω 是单连通的, 这个条件对于解的存在性也是充分的. 要注意下面证明与经典 Poincaré 引理 (定理 6.17-2) 证明的类似性.

定理 6.20-1 (Pfaff 方程组解的存在性和唯一性[100]) 设 Ω 是 $\mathbb{R}^N$ 的一个单连通

[99] 冠名源自 Johann Friedrich Pfaff (1765—1825), 他应该感到荣幸的是, Carl Friedrich Gauss 是他的博士生之一.

[100] 这一结果可以追溯到:

E. CARTAN [1927]: Sur la possibilité de plonger un espace riemannien donné dans un espace euclidien. Annales de la Société Polonaise de Mathématiques **6**, 1–7.

后来这一结果被拓展到非线性 Pfaff 方程组, 见:

T. Y. THOMAS [1934]: Systems of total differential equations defined over simply connected domains. Annals of Mathematics **35**, 730–734.

开子集, 给定矩阵场 $\boldsymbol{\Gamma}_i \in \mathcal{C}^1(\Omega; \mathbb{M}^n), 1 \leqslant i \leqslant N$, 它们满足

$$\partial_i \boldsymbol{\Gamma}_k(x) - \partial_k \boldsymbol{\Gamma}_i(x) + \boldsymbol{\Gamma}_i(x)\boldsymbol{\Gamma}_k(x) - \boldsymbol{\Gamma}_k(x)\boldsymbol{\Gamma}_i(x) = \mathbf{0}, \quad x \in \Omega, 1 \leqslant i, k \leqslant N.$$

又设给定一点 $x^0 \in \Omega$ 及一矩阵 $\boldsymbol{F}^0 \in \mathbb{M}^n$. 则存在一个, 且只有一个矩阵场 $\boldsymbol{F} \in \mathcal{C}^2(\Omega; \mathbb{M}^n)$ 满足

$$\partial_i \boldsymbol{F}(x) = \boldsymbol{F}(x)\boldsymbol{\Gamma}_i(x), \quad x \in \Omega, 1 \leqslant i \leqslant N,$$
$$\boldsymbol{F}(x^0) = \boldsymbol{F}^0.$$

证明 指定 A_{ij} 表示任一矩阵 $\boldsymbol{A} \in \mathbb{M}^n$ 的第 i 行, 第 j 列的元素. 特定的符号 Γ_{ij}^p 表示一矩阵 $\boldsymbol{\Gamma}_i \in \mathbb{M}^n$ 的第 p 行, 第 j 列的元素, 其中 i 是在 $\{1, \cdots, N\}$ 范围内变化取值的整数. 在此仍用关于重复的下标或上标的求和约定, 上、下标则在 $\{1, \cdots, N\}$ 或 $\{1, \cdots, n\}$ 中变化取值.

(i) 设 x^1 是集合 Ω 中与 x^0 不同的任一点. 因为 Ω 特别地还是弧连通的, 故存在一个在 Ω 中联结 x^0 与 x^1 的路径 $\gamma = (\gamma^i) \in \mathcal{C}^1([0,1]; \mathbb{R}^N)$; 这意味着

$$\gamma(0) = x^0, \gamma(1) = x^1, \text{ 及 } \gamma(t) \in \Omega, \quad \text{对所有 } 0 \leqslant t \leqslant 1.$$

假设一矩阵场 $\boldsymbol{F} = (F_{\ell j}) \in \mathcal{C}^1(\Omega; \mathbb{M}^n)$ 满足

$$\partial_i \boldsymbol{F}(x) = \boldsymbol{F}(x)\boldsymbol{\Gamma}_i(x) \quad \text{或等价地}$$
$$\partial_i F_{\ell j}(x) = \Gamma_{ij}^p(x) F_{\ell p}(x) \quad \text{在每一点 } x \in \Omega.$$

则对每个整数 $1 \leqslant \ell \leqslant n$, 由下式

$$\zeta_j(t) := F_{\ell j}(\gamma(t)), \quad 0 \leqslant t \leqslant 1, 1 \leqslant j \leqslant n$$

定义的 n 个函数 $\zeta_j \in \mathcal{C}^1([0,1])$ (这里为简单起见, 省略了其对于 ℓ 的依赖性) 满足下述由 n 个未知函数的 n 个常微分方程组成的线性方程组的 Cauchy 问题:

$$\frac{d\zeta_j}{dt}(t) = \Gamma_{ij}^p(\gamma(t)) \frac{d\gamma^i}{dt}(t) \zeta_p(t), \quad 0 \leqslant t \leqslant 1,$$
$$\zeta_j(0) = \zeta_j^0,$$

其中初始值 ζ_j^0 由

$$\zeta_j^0 := F_{\ell j}^0$$

给出 (注意, 这些 Cauchy 问题的差别只是其初值 ζ_j^0 不同).

因为形如 (其符号不说自明)

$$\frac{d\boldsymbol{\zeta}(t)}{dt} = \boldsymbol{A}(t)\boldsymbol{\zeta}(t), \quad 0 \leqslant t \leqslant 1,$$
$$\boldsymbol{\zeta}(0) = \boldsymbol{\zeta}^0$$

的 Cauchy 问题有且只有一个解 $\boldsymbol{\zeta} \in \mathcal{C}^1([0,1],\mathbb{R}^n)$ 如果 $\boldsymbol{A} \in \mathcal{C}([0,1];\mathbb{M}^n)$ (定理 3.8-2), 所以前述 Cauchy 问题中每一个均有唯一解.

这个结果顺便也证明了, 如果未知的场 $\boldsymbol{F} = (F_{\ell j})$ 存在, 它是唯一的.

(ii) 对于给定的整数 $\ell \in \{1, \cdots, n\}$ 为了使得解上述 Cauchy 问题所得到的值 $\boldsymbol{\zeta}_j(1)$ 可以成为可接受的未知量值 $F_{\ell j}(x^1)$, 它们当然必须与联结 x^0 与 x^1 的路径的选取无关.

设 $\gamma_0 \in \mathcal{C}^1([0,1];\mathbb{R}^N)$ 和 $\gamma_1 \in \mathcal{C}^1([0,1];\mathbb{R}^N)$ 是在 Ω 内联结 x^0 与 x^1 的两个路径. 开集 Ω 是单连通的, 故存在 Ω 中联结 γ_0 与 γ_1 的同伦 $\boldsymbol{G} = (\boldsymbol{G}^i) : [0,1] \times [0,1] \to \mathbb{R}^N$ (定理 6.17-1), 即使得

$$
\begin{aligned}
&\boldsymbol{G}(\cdot,0) = \gamma_0, \quad \boldsymbol{G}(\cdot,1) = \gamma_1,\\
&\boldsymbol{G}(t,\lambda) \in \Omega \quad \text{对所有 } 0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1,\\
&\boldsymbol{G}(0,\lambda) = x^0 \quad \text{及 } \boldsymbol{G}(1,\lambda) = x^1 \text{ 以所有 } 0 \leqslant \lambda \leqslant 1,
\end{aligned}
$$

而且在下述意义下足够光滑

$$
\begin{gathered}
\boldsymbol{G} \in \mathcal{C}^1([0,1] \times [0,1];\mathbb{R}^N) \quad \text{及}\\
\frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{G}}{\partial \lambda}\right) = \frac{\partial}{\partial \lambda}\left(\frac{\partial \boldsymbol{G}}{\partial t}\right) \in \mathcal{C}([0,1] \times [0,1];\mathbb{R}^N).
\end{gathered}
$$

设对每一个 $0 \leqslant \lambda \leqslant 1, \boldsymbol{\zeta}(\cdot,\lambda) = (\zeta_j(\cdot,\lambda)) \in \mathcal{C}^1([0,1];\mathbb{R}^N)$ 表示相应于联结 x^0 与 x^1 的路径 $\boldsymbol{G}(\cdot,\lambda)$ 的 Cauchy 问题的解. 我们有

$$
\begin{aligned}
&\frac{\partial \zeta_j}{\partial t}(t,\lambda) = \Gamma^p_{ij}(\boldsymbol{G}(t,\lambda))\frac{\partial G^i}{\partial t}(t,\lambda)\zeta_p(t,\lambda) \quad \text{对所有 } 0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1,\\
&\zeta_j(0,\lambda) = \zeta^0_j \quad \text{对所有 } 0 \leqslant \lambda \leqslant 1.
\end{aligned}
$$

我们的目标是要证明

$$
\frac{\partial \zeta_j}{\partial \lambda}(1,\lambda) = 0 \quad \text{对所有 } 0 \leqslant \lambda \leqslant 1,
$$

因为这个式子意味着 $\zeta_j(1,0) = \zeta_j(1,1)$, 正是我们希望得到的结论. 为此目的, 直接求导得, 对所有 $0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1$ 有

$$
\frac{\partial}{\partial \lambda}\left(\frac{\partial \zeta_j}{\partial t}\right) = \{\Gamma^q_{ij}\Gamma^p_{kq} + \partial_k\Gamma^p_{ij}\}\zeta_p\frac{\partial G^k}{\partial \lambda}\frac{\partial G^i}{\partial t} + \Gamma^p_{ij}\zeta_p\frac{\partial}{\partial \lambda}\left(\frac{\partial G^i}{\partial t}\right) + \sigma_q\Gamma^q_{ij}\frac{\partial G^i}{\partial t},
$$

其中

$$
\sigma_j := \frac{\partial \zeta_j}{\partial \lambda} - \Gamma^p_{kj}\zeta_p\frac{\partial G^k}{\partial \lambda},
$$

这是一方面 (在上述以及以下各式中, 以 $\Gamma^q_{ij}, \partial_k\Gamma^p_{ij}$ 等式表示 $\Gamma^q_{ij}(\boldsymbol{G}(\cdot,\cdot)), \partial_k\Gamma^p_{ij}(\boldsymbol{G}(\cdot,\cdot))$ 等).

另一方面, 直接对定义函数 σ_j 的方程求导可得, 对所有 $0 \leqslant t \leqslant 1, 0 \leqslant \lambda \leqslant 1$ 有

$$\frac{\partial}{\partial t}\left(\frac{\partial \zeta_j}{\partial \lambda}\right) = \frac{\partial \sigma_j}{\partial t} + \left\{\partial_i \Gamma^p_{kj} \frac{\partial G^i}{\partial t}\zeta_p + \Gamma^q_{kj}\frac{\partial \zeta_q}{\partial t}\right\}\frac{\partial G^k}{\partial \lambda} + \Gamma^p_{ij}\zeta_p \frac{\partial}{\partial t}\left(\frac{\partial G^i}{\partial \lambda}\right).$$

但 $\dfrac{\partial \zeta_j}{\partial t} = \Gamma^p_{ij}\dfrac{\partial G_i}{\partial t}\zeta_p$, 所以我们也有

$$\frac{\partial}{\partial t}\left(\frac{\partial \zeta_j}{\partial \lambda}\right) = \frac{\partial \sigma_j}{\partial t} + \{\partial_i \Gamma^p_{kj} + \Gamma^q_{kj}\Gamma^p_{iq}\}\zeta_p \frac{\partial G^i}{\partial t}\frac{\partial G^k}{\partial \lambda} + \Gamma^p_{ij}\zeta_p \frac{\partial}{\partial t}\left(\frac{\partial G^i}{\partial \lambda}\right).$$

所以, 将上面两个关系式相减, 并注意到根据假设有 $\dfrac{\partial}{\partial \lambda}\left(\dfrac{\partial \zeta_j}{\partial t}\right) = \dfrac{\partial}{\partial t}\left(\dfrac{\partial \zeta_j}{\partial \lambda}\right)$ 和 $\dfrac{\partial}{\partial \lambda}\left(\dfrac{\partial G^i}{\partial t}\right) = \dfrac{\partial}{\partial t}\left(\dfrac{\partial G^i}{\partial \lambda}\right)$, 我们推得

$$\frac{\partial \sigma_j}{\partial t} + \{\partial_i \Gamma^p_{kj} - \partial_k \Gamma^p_{ij} + \Gamma^q_{kj}\Gamma^p_{iq} - \Gamma^q_{ij}\Gamma^p_{kq}\}\zeta_p \frac{\partial G^k}{\partial \lambda}\frac{\partial G^i}{\partial t} - \Gamma^q_{ij}\frac{\partial G^i}{\partial t}\sigma_q = 0.$$

但是 $\{\partial_i \Gamma^p_{kj} - \partial_k \Gamma^p_{ij} + \Gamma^q_{kj}\Gamma^p_{iq} - \Gamma^q_{ij}\Gamma^p_{kq}\}$ 正是矩阵场 $\partial_i \boldsymbol{\Gamma}_k - \partial_k \boldsymbol{\Gamma}_i + \boldsymbol{\Gamma}_i \boldsymbol{\Gamma}_k - \boldsymbol{\Gamma}_k \boldsymbol{\Gamma}_i$ 的第 p 行、第 j 列的元素, 根据假设其为零. 所以

$$\partial_i \Gamma^p_{kj} - \partial_k \Gamma^p_{ij} + \Gamma^q_{kj}\Gamma^p_{iq} - \Gamma^q_{ij}\Gamma^p_{kq} = 0,$$

这是一方面. 另一方面

$$\sigma_j(0,\lambda) = \frac{\partial \zeta_j}{\partial \lambda}(0,\lambda) - \Gamma^p_{kj}(\boldsymbol{G}(0,\lambda))\zeta_p(0,\lambda)\frac{\partial G^k}{\partial \lambda}(0,\lambda) = 0,$$

这是因为 $\zeta^0_j(0,\lambda) = \zeta^0_j$ 及 $\boldsymbol{G}(0,\lambda) = x^0$ 对所有 $0 \leqslant \lambda \leqslant 1$. 所以, 对参数 $\lambda \in [0,1]$ 的每一固定值, 每个函数 $\sigma_j(\cdot,\lambda)$ 都满足一个常微分方程的 Cauchy 问题, 即

$$\frac{d\sigma_j}{dt}(t,\lambda) = \Gamma^q_{ij}(\boldsymbol{G}(t,\lambda))\frac{\partial G^i}{\partial t}(t,\lambda)\sigma_q(t,\lambda), \quad 0 \leqslant t \leqslant 1, \quad \sigma_j(0,\lambda) = 0.$$

但这种 Cauchy 问题的解是唯一的 (定理 3.8-2); 因此 $\sigma_j(t,\lambda) = 0$ 对所有 $0 \leqslant t \leqslant 1$. 特别地, 有

$$\sigma_j(1,\lambda) = \frac{\partial \zeta_j}{\partial \lambda}(1,\lambda) - \Gamma^p_{kj}(\boldsymbol{G}(1,\lambda))\zeta_p(1,\lambda)\frac{\partial G^k}{\partial \lambda}(1,\lambda) = 0, \quad \text{对所有 } 0 \leqslant \lambda \leqslant 1.$$

又因为 $\boldsymbol{G}(1,\lambda) = x^1$ 对所有 $0 \leqslant \lambda \leqslant 1$, 故 $\dfrac{\partial \zeta_j}{\partial \lambda}(1,\lambda) = 0$ 对所有 $0 \leqslant \lambda \leqslant 1$.

对于每个整数 $\ell \in \{1,\cdots,n\}$, 我们现在可以毫不含糊地定义向量场 $(F_{\ell j}) : \Omega \to \mathbb{R}^n$, 令:

$$F_{\ell j}(x^1) := \zeta_j(1) \quad \text{对任意 } x^1 \in \Omega,$$

其中 $\gamma \in \mathcal{C}^1([0,1];\mathbb{R}^N)$ 是 Ω 中联结 x^0 与 x^1 的任一路径, 而向量场 $(\zeta_j) \in \mathcal{C}^1([0,1])$ 是相应于这样一个路径的 Cauchy 问题:

$$\begin{aligned} &\frac{d\zeta_j}{dt}(t) = \Gamma^p_{ij}(\gamma(t))\frac{d\gamma^i}{dt}(t)\zeta_p(t), \quad 0 \leqslant t \leqslant 1, \\ &\zeta_j(0) = \zeta^0_j \end{aligned}$$

的解.

(iii) 为了说明这样一个向量场确实是我们要寻求的未知矩阵场的第 ℓ 个行向量场, 我们需证明, 对于在上述 Cauchy 问题中固定的 $\ell, (F_{\ell j})_{j=1}^n \in \mathcal{C}^1(\Omega;\mathbb{R}^n)$ 而且这个向量场满足偏微分方程组 $\partial_i F_{\ell j} = \Gamma_{ij}^p F_{\ell p}$ 在 Ω 中.

在下文中, 设 x 是 Ω 中任一点, 而整数 $i \in \{1,\cdots,n\}$ 是固定的. 则存在 $x^1 \in \Omega$, 一个联结 x^0 及 x^1 的路径 $\gamma \in \mathcal{C}^1([0,1];\mathbb{R}^n), \tau \in]0,1[$ 以及包含 τ 的开区间 $I \subset [0,1]$ 使得

$$\gamma(t) = x + (t-\tau)\boldsymbol{e}_i \quad \text{对 } t \in I,$$

其中 $\boldsymbol{e}_i$ 是 $\mathbb{R}^N$ 中第 i 个基向量. 因为每个函数 ζ_j 都是在 $[0,1]$ 中连续可微的而且满足 $\dfrac{d\zeta_j}{dt}(t) = \Gamma_{kj}^p(\gamma(t))\dfrac{d\gamma^k}{dt}(t)\zeta_p(t)$ 对所有 $0 \leqslant t \leqslant 1$, 又因为 $\dfrac{d\gamma^k}{dt}(\tau) = \delta_i^k$, 我们有

$$\begin{aligned}\zeta_j(t) &= \zeta_j(\tau) + (t-\tau)\frac{d\zeta_j}{dt}(\tau) + o(t-\tau)\\ &= \zeta_j(\tau) + (t-\tau)\Gamma_{ij}^p(\gamma(\tau))\zeta_p(\tau) + o(t-\tau)\end{aligned}$$

对所有 $t \in I$. 等价地, 即

$$F_{\ell j}(x + (t-\tau)\boldsymbol{e}_i) = F_{\ell j}(x) + (t-\tau)\Gamma_{ij}^p(x)F_{\ell p}(x) + o(t-x).$$

这个式子说明, 每个函数 $F_{\ell j}$ 在集合 Ω 中具有偏导数, 在每个 $x \in \Omega$, 这些偏导数由下式给出:

$$\partial_i F_{\ell n}(x) = \Gamma_{ij}^p(x)F_{\ell p}(x),$$

或矩阵形式 $\partial_i \boldsymbol{F}(x) = \boldsymbol{F}(x)\boldsymbol{\Gamma}_i(x)$.

(iv) 从 (iii) 我们知道, 矩阵场 $(F_{\ell j})$ 在 Ω 中是 $\mathcal{C}^1$ 类的 (其偏导数在 Ω 中连续) 而且在 Ω 中满足偏微分方程组 $\partial_i F_{\ell j} = \Gamma_{ij}^p F_{\ell p}$. 微分这些方程就说明矩阵场 $(F_{\ell j})$ 在 Ω 中实际上是 $\mathcal{C}^2$ 类的. 这就完成了证明. □

关于矩阵场 $\boldsymbol{\Gamma}_i : \Omega \to \mathbb{M}^n, 1 \leqslant i \leqslant N$, 的正则性假设可以适当减弱. 更明确地说, 如果 $\boldsymbol{\Gamma}_i \in \mathcal{C}(\Omega;\mathbb{M}^n), 1 \leqslant i \leqslant N$, 定理 6.20-1 中关于 Pfaff 方程组解的存在性仍然成立, 只是在这种情况下解 $\boldsymbol{F}$ 属于空间 $\mathcal{C}^1(\Omega,\mathbb{M}^n)$[101]; 如果 Ω 是 $\mathbb{R}^N$ 中的区域, 而 $\Gamma_i \in L^p(\Omega;\mathbb{M}^n)$ 对某 $p > n$, 在这种情况下解 $\boldsymbol{F}$ 属于空间 $W^{1,p}(\Omega;\mathbb{M}^n)$[102]. 在这些情况下, 矩阵场 $\boldsymbol{\Gamma}_i$ 应满足的相容性条件自然地要理解为在分布意义下成立.

---

[101] P. Hartman; A. Wintner [1950]: On the fundamental equations of differential geometry. American Journal of Mathematics **72**, 757–774.

[102] S. Mardare [2005]: On Pfaff systems with L^p coefficients and their applications in differentical geometry. Journal de Mathématiques Pures et Appliquées **84**, 1659–1692.

S. Mardare [2007]: On systems of first order linear partial differential equations with L^p coefficients. Advances in Differential Equations **12**, 301–306.

习题

6.20-1 设 Ω 是 $\mathbb{R}^N$ 的单连通开子集, 又设给定矩阵场 $\boldsymbol{A}_i \in \mathcal{C}^1(\Omega; \mathbb{M}^n), 1 \leqslant i \leqslant k, \boldsymbol{B}_j \in \mathcal{C}^1(\Omega; \mathbb{M}^m), 1 \leqslant j \leqslant N$, 及 $\boldsymbol{C}_k \in \mathcal{C}^1(\Omega; \mathbb{M}^{m\times n}), 1 \leqslant k \leqslant N$, 它们满足

$$\partial_i \boldsymbol{A}_j - \partial_j \boldsymbol{A}_i + \boldsymbol{A}_i\boldsymbol{A}_j - \boldsymbol{A}_j\boldsymbol{A}_i = \boldsymbol{0} \quad \text{在 } \Omega \text{ 中},$$

$$\partial_i \boldsymbol{B}_j - \partial_j \boldsymbol{B}_i + \boldsymbol{B}_j\boldsymbol{B}_i - \boldsymbol{B}_i\boldsymbol{B}_j = \boldsymbol{0} \quad \text{在 } \Omega \text{ 中},$$

$$\partial_i \boldsymbol{C}_j - \partial_j \boldsymbol{C}_i + \boldsymbol{C}_i\boldsymbol{A}_j - \boldsymbol{C}_j\boldsymbol{A}_i + \boldsymbol{B}_j\boldsymbol{C}_i - \boldsymbol{B}_i\boldsymbol{C}_j = \boldsymbol{0} \quad \text{在 } \Omega \text{ 中},$$

另设点 $x^0 \in \Omega$ 和矩阵 $\boldsymbol{F}^0 \in \mathbb{M}^{m\times m}$ 是给定的. 证明, 存在一个且只有一个矩阵场 $\boldsymbol{F} \in \mathcal{C}^2(\Omega; \mathbb{M}^{m\times n})$, 它是下述 Pfaff 方程组的解:

$$\partial_i \boldsymbol{F} = \boldsymbol{F}\boldsymbol{A}_i + \boldsymbol{B}_i\boldsymbol{F} + \boldsymbol{C}_i \quad \text{在 } \Omega \text{ 中},$$

$$\boldsymbol{F}(x^0) = \boldsymbol{F}^0.$$

注 这个问题对定理 6.20-1 中讨论的 Pfaff 方程组给出了一种推广, 定理 6.20-1 相应这里 $\boldsymbol{B}_j = \boldsymbol{0}$ 及 $\boldsymbol{C}_k = \boldsymbol{0}$ 的特殊情况; 这个问题同样也是经典 Poincaré 引理 (定理 6.17-2) 的推广, 后者相应于这里 $m = n = 1$ 和 $\boldsymbol{A}_i = \boldsymbol{0}$ 及 $\boldsymbol{B}_j = \boldsymbol{0}$ 的特殊情况.

文 献 注 释

下面文献注释中列出的书籍及参考性论文只是为读者提供可选取的文献, 作为本书有益的补充; 但在此并不想做得详尽无遗. 原始的参考文献均散布在各章的脚注中.

第 1 章: 实分析和函数论

关于详细的证明、补充, 以及其他的参考文献, 建议读者参阅经典教科书, 如 Dieudonné [1960], Royden [1963], Hewiti & Stromberg [1965] 或 Rudin [1966], 在这些书中, 对实分析的核心课题给予详尽的讨论. 更新的论述有: Schwartz [1970, 1991], Folland [1984], Lang [1993], DiBenedetto [2002], Krantz [2004], Jost [2005], Knapp [2005a, 2005b] 或 Amann & Escher [2005, 2008, 2009]. Li [2011] 的书包含许多有相当难度的实分析习题.

关于集合论更专门且深入的讨论可参阅 Cohen [1966], Bourbaki [1970] 或 Halmos [1982]; 关于一般拓扑的可见 Kelley [1955], Taylor [1965], Choquet [1966] 或 Bourbaki [1966a, 1966b]; 关于 Lebesgue 测度和积分的可见 Halmos [1950], Munroe [1953], Schwartz [1993a, 1993b], Rana [2002] 或 Benedetto & Czaja [2009].

函数论的深入分析可参阅 Stein [1970], Evans & Gariepy [1992], Carothers [2000] 或 Stein & Shakarchi [2005]. 关于 $\mathbb{R}^n$ 中区域和 Green 公式透彻的处理在 Nečas [1967] 中给出.

第 2 到第 5 章: 赋范向量空间; Banach 空间; 内积空间和 Hilbert 空间; 线性泛函分析中的重要定理

这些章节的内容含盖了线性泛函分析的基本结果, 下述经典教科书可作为其有益的补充: BANACH [1932] (这部专著奠定了现代线性泛函分析的基础), Riesz & NAGY [1955], TAYLOR [1958] (后来修订并扩充为 TAYLOR & LAY [1980]), DUNFORD & SCHWARTZ [1958, 1963, 1971] 不朽的专论, GOFFMAN & PEDRICK [1965], YOSIDA [1966] (经典中的经典), RUDIN [1973], SCHECHTER [1971], HALMOS [1974], DISTEL [1975], KREYSZIG [1978], KESAVAN [1989], CONWAY [1990], ZEIDLER [1995a, 1995b], DEBNATH & MIKUSIŃSKI [1999], AUBIN [2000], LAX [2002], HA [2007], KESAVAN [2009], ODEN & DEMKOWICZ [2010], STEIN & SHAKARCHI [2011] 或 BREZIS [2011] (极成功的原法文版教科书 BREZIS [1983] 的英译扩充版).

谱问题的泛函分析 (除紧自伴算子外, 本书未予以讨论) 可参阅 DUNFORD & SCHWARTZ [1971], FRIEDRICHS [1981] 或 DAUTRAY & LIONS [2000c].

将线性泛函分析与插值论, 逼近论或数值分析相结合的讨论可在下述著作中找到: DAVIS [1963], CHENEY [1966], DAVIS & RABINOWITZ [1975], CROUZEIX & MIGNOT [1983], CHATELIN [1983], ATKINSON & HAN [2009] 或 MHASKAR & PAI [2007].

关于矩阵论, 实际上是有限维空间中的泛函分析, 其相关内容可在下述著作中找到: GANTMACHER [1959], VARGA [1962], HOUSEHOLDER [1964], STRANG [1976, 2009], HORN & JOHNSON [1985, 1991], CIARLET [1987], SERRE [2010].

对泛函分析的历史以及对本书中引用过的泛函分析学家的自传、传记和照片感兴趣的读者, 可查阅以下著作, 乐于细读或只是粗略地浏览均宜: REID [1970, 1976], BOURBAKI [1974], ULAM [1976], WESTFALL [1980], DIEUDONNÉ [1981], MAUDLIN [1981], ALBERS & ALEXANDERSON [1985], HALMOS [1985, 1987], PÓLYA [1987], CURIEN & SCHMIDT [1990], RUDIN [1997], BATTERSON [2000], SCHWARTZ [2001], MAZ'YA & SHAPOSHNIKOVA [1998], PIETSCH [2007], JAKIMOWICZ & MIRANOVICZ [2011] 或 GRAY [2012] (在此只列出这类书籍的很少一部分).

第 6 章: 线性偏微分方程

关于空间 $\mathcal{D}(\Omega)$ 和 $\mathcal{D}'(\Omega)$ 拓扑的更多细节, 参阅 YOSIDA [1966], VO-KHAC [1972a, 1972b], HÖRMANDER [1983, 第 1—7 章], DUISTERMAAT & KOLK [2010], 当然还有 SCHWARTZ [1966] 驰名学界的专著, 他创立了分布论并建立了系统的理论. 关于这个理论的具有启发性的介绍给出在 SCHWARTZ [1965] 中.

关于 Sobolev 空间详尽的讨论, 对 Hilbert 空间情况, 见 LIONS & MAGENES [1972, 第 1 章] 和 DAUTRAY & LIONS [2000b, 第 4 章]; 对一般情况, 见 LIONS [1965, 第

1—3 章], Nečas [1967, 第 2 章], Adams [1975], Attouch, Buttazzo & Michaille [2006], Tartar [2007] 或 Brezis [2011, 第 9 章].

由于关于线性偏微分方程的教科书数量庞大, 我们将引用目标局限在少量经挑选的教科书, 它们的内容及处理问题的方法 (至少部分地) 类似于本章中所给出的 (本章后几节讨论的更专门的课题, 如 Poincaré 引理、Saint-Venant 引理、Donati 引理以及 Pfaff 方程组等, 其参考文献已经在脚注中给出).

更具体地说, Buttazzo, Giaquita & Hildebrandt [1998], Chipot [2009] 或 Cioranescu, Donato & Roque [2012] 都是优秀的入门书. 在更高的水平上, 读者可参阅如下经典教科书, 如 Nečas [1967], Kinderlehrer & Stampacchia [1980], Gilbarg & Trudinger [1988], Taylor [1996a, 1996b], Stakgold [1998], Attouch, Buttazzo & Michaille [2006], Sauvigny [2006a, 2006b], Evans [2010], DiBenedetto [2010], 而尤其是专著 Dautray & Lions [2000a, 2000b, 2000c, 2000d, 2000e, 2000f], 其中详尽讨论了范围极广泛的各种应用.

奇性在 Grisvard [1992] 中给予详尽地讨论. 具有 "小" 参数的变分问题, 如奇摄动问题或均匀化问题, 其数学分析实质上是 Lions [1973] 开创的. 更新的讨论可参阅 Cioranescu & Donato [1999], Cioranescu & Saint Jeint Paulin [1999], 或 Ciarlet [1997, 2000] 关于对线性化板和壳理论的应用. 对于特定的椭圆问题的渐近分析已在 Chipot [2002], Tartar [2009] 或 Ghergu & Rădulescu [2008] 中予以讨论.

产生于线性化弹性的问题的变分形式 (包括模型为变分不等式的情况) 在 Duvaut & Lions [1972], Fichera [1972a, 1972b] 和 Nečas & Hlaváček [1981] 中给予了详尽地讨论.

由变分方程或变分不等式表示的模型问题解的逼近在已在下述著作中给予充分的分析讨论: Ciarlet [1978, 1991], Glowinski, Lions & Trémolières [1981], Raviart & Thomas [1983], Glowinski [1984], Girault & Raviart [1986], Brezzi & Fortin [1991], Babuška & Osborn [1991], Roberts & Thomas [1991] 或 Brenner & Scott [2002] (这里列出的文献远不是详尽无遗的).

参考文献

R. ABRAHAM; J.E. MARSDEN; T. RATIU [1988]: *Manifolds, Tensor Analysis, and Applications, Second Edition*, Springer, New York (First Edition: 1983, Addison-Wesley).

R.A. ADAMS [1975]: *Sobolev Spaces*, Academic Press, New York.

N.I. AKHIEZER; I.M. GLAZMAN [1961]: *Theory of Linear Operators in Hilbert Spaces, Volume* 1, Ungar, New York.

J.L. AKIAN [2003]: A simple proof of the ellipticity of Koiter's model, *Analysis and Applications* **1**, 1–16.

D.J. ALBERS; G.L. ALEXANDERSON, editors [1985]: *Mathematical People—Profiles and Interieurs*, Birkhäuser, Boston.

H. AMANN [1976]: Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces, *SIAM Review* **18**, 620–709.

H. AMANN; J. ESCHER [2005]: *Analysis* I, Birkhäuser, Boston (translation of the original German edition, *Analysis* I, Birkhäuser, Basel, 1998).

H. AMANN; J. ESCHER [2008]: *Analysis* II, Birkhäuser, Boston (translation of the original German edition, *Analysis* II, Birkhäuser, Basel, 1999).

H. AMANN; J. ESCHER [2009]: *Analysis* III, Birkhäuser, Boston (translation of the original German edition, *Analysis* III, Birkhäuser, Basel, 2001).

C. AMROUCHE; P.G. CIARLET; L. GRATIE; S. KESAVAN [2006]: On the characterization of matrix fields as linearized strain tensor fields, *Journal de Mathématiques Pures et Appliquées* **86**, 116–132.

C. AMROUCHE; V. GIRAULT [1994]: Decomposition of vector spaces and application to the Stokes problem in arbitrary dimension, *Czechoslovak Mathematical Journal* **44**, 109–140.

S.S. ANTMAN [1970]: Existence of solutions of the equilibrium equations for nonlinearly elastic rings and arches, *Indiana University Mathematics Journal* **20**, 281–302.

S.S. ANTMAN [1983]: Regular and singular problems for large elastic deformations of tubes, wedges, and cylinders, *Archive for Rational Mechanics and Analysis* **82**, 1–52.

S.S. ANTMAN [2005]: *Nonlinear Problems of Elasticity*, Springer, Berlin (First Edition: 1995).

D.N. ARNOLD; R.S. FALK; R. WINTHER [2006]: Finite element exterior calculus, homological techniques, and applications, in *Acta Numerica, Volume* 15 (A. ISERLES, editor), pp. 1-155, Cambridge University Press, Cambridge, UK.

N. ARONSZAJN; K.T. SMITH [1957]: Characterization of positive reproducing kernels. Applications to Green's functions, *American Journal of Mathematics* **79**, 611–622.

C. ARZELÀ [1883]: Un' osservazione intorno alle serie di funzioni, *Rendiconti delle Sessioni dell' Accademia Reale delle Scienze dell' Istituto di Bologna*, 142–159.

C. ASCOLI [1883]: Le curve limiti di una varietà data di curve, *Atti della Accademia Nazionale dei Lincei, Classe di Scienze Fisiche, Matematiche e Naturali* **18**, 521–586.

K.E. ATKINSON; W. HAN [2009]: *Theoretical Numerical Analysis: A Functional Analysis Framework, Third Edition*, Springer, New York (First Edition: 2001).

H. ATTOUCH [1984]: *Variational Convergence for Functions and Operators*, Pitman, Boston.

H. ATTOUCH; G. BUTTAZZO; G. MICHAILLE [2006]: *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*, SIAM, Philadelphia.

J.P. AUBIN [1993]: *Optima and Equilibria—An Introduction to Nonlinear Analysis*, Springer, Berlin.

J.P. AUBIN [2000]: *Applied Functional Analysis, Second Edition*, Wiley-Interscience, New York (First Edition: 1979).

I. BABUŠKA [1971]: Error bound for finite element method, *Numerische Mathematik* **16**, 322–333.

I. BABUŠKA; A.K. AZIZ [1976]: On the angle condition in the finite element method, *SIAM Journal on Numerical Analysis* **13**, 214–226.

I. BABUŠKA; J. OSBORN [1991]: Eigenvalue problems, in *Handbook of Numerical Analysis, Volume* II (P.G. CIARLET & J.L. LIONS, editors), pp. 641–784, North-Holland, Amsterdam.

R. BAIRE [1899]: Sur les fonctions de variables réelles, *Annali di Matematica Pura ed Applicata* **3**, 1–123.

J. BALL [1977]: Convexity conditions and existence theorems in nonlinear elasticity, *Archive for Rational Mechanics and Analysis* **63**, 337–403.

J. BALL [1981]: Global invertibility of Sobolev functions and the interpenetration of matter, *Proceedings of the Royal Society, Edinburgh* **88A**, 315–328.

J.M. BALL; R.J. KNOPS; J.E. MARSDEN [1978]: Two examples in nonlinear elasticity, in *Proceedings-Conference in Nonlinear Analysis, Besançon*, pp. 41–49, Lecture Notes in Mathematics, Volume 466, Springer, Berlin.

S. BANACH [1922]: Sur les opérations dans les ensembles abstraits et leurs applications aux

équations intégrales, *Fundamenta Mathematicae* **3**, 133–181.

S. BANACH [1932]: *Théorie des Opérations Linéaires*, Monografje Matematyczne, Volume 1, Warsaw.

S. BANACH; S. SAKS [1930]: Sur la convergence forte dans le champ L^p, *Studia Mathematica* **2**, 51–57.

S. BANACH; H. STEINHAUS [1927]: Sur le principe de la condensation de singularités, *Fundamenta Mathematicae* **9**, 50–61.

S. BATTERSON [2000]: *Stephen Smale: The Mathematician Who Broke the Dimension Barrier*, American Mathematical Society, Providence, RI.

R. BEALS; R. WONG [2010]: *Special Functions: A Graduate Text*, Cambridge University Press, Cambridge, UK.

P.R. BEESACK; E. HUGHES; M. ORTEL [1979]: Rotund complex linear spaces, *Proceedings of the American Mathematical Society* **75**, 42–44.

J.J. BENEDETTO; W. CZAJA [2009]: *Integration and Modern Analysis*, Birkhäuser, Boston.

A. BEN-ISRAEL; T.N.E. GREVILLE [2003]: *Generalized Inverses: Theory and Applications*, Second Edition, Springer.

M.S. BERGER [1967]: On the von Kármán equations and the buckling of a thin elastic plate. I. The clamped plate, *Communications on Pure and Applied Mathematics* **20**, 687–719.

M.S. BERGER [1977]: *Nonlinearity and Functional Analysis*, Academic Press, New York.

M. BERGER [2003]: *A Panoramic View of Riemannian Geometry*, Springer, Berlin.

M. BERGER; B. GOSTIAUX [1987]: *Géométrie Différentielle: Variétés, Courbes et Surfaces*, Presses Universitaires de France, Paris.

S. BERGMAN; M. SCHIFFER [1948]: Kernel functions in the theory of partial differential equations of elliptic type, *Duke Mathematical Journal* **15**, 535–566.

A. BERMAN; R.J. PLEMMONS [1994]: *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics, Vol. 9, SIAM, Philadelphia.

M. BERNADOU; P.G. CIARLET [1976]: Sur l'ellipticité du modèle linéaire de coques de W.T. Koiter, in *Computing Methods in Applied Sciences and Engineering* (R. GLOWINSKI & J.L. LIONS, editors), pp. 89–136, Lecture Notes in Economics and Mathematical Systems, **134**, Springer, Heidelberg.

M. BERNADOU; P.G. CIARLET; B. MIARA [1994]: Existence theorems for two-dimensional linear shell theories, *Journal of Elasticity* **34**, 111–138.

J.M.E. BERNARD [2011]: Density results in Sobolev spaces whose elements vanish on a part of the boundary, *Chinese Annals of Mathematics, Series B*, **32**, 823–846.

S.N. BERNSTEIN [1912]: Démonstration du théorème de Weierstrass fondée sur le calcul de probabilités, *Communications of the Kharkov Mathematical Society* **13**, 1–2.

S.N. BERNSTEIN [1932]: Complément à l'article de E. Voronovskaya "Détermination de la forme asymptotique de l'approximation des fonctions par les polynômes de M. Bernstein,"

Doklady Akademii Nauk SSSR **4**, 86–92.

G. BIRKHOFF [1946]: Tres observaciones sobre el algebra lineal, *Universidad Nacional de Tucumán Revista A* **5**, 147–151.

E. BISHOP; R.R. PHELPS [1961]: A proof that every Banach space is subreflexive, *Bulletin of the American Mathematical Society* **67**, 97–98.

A. BLOUZA; H. LE DRET: [1999]: Existence and uniqueness for the linear Koiter model for shells with little regularity, *Quarterly of Applied Mathematics* **57**, 317–337.

A.B. BOGHOSSIAN; P.D. JOHNSON, JR. [1990]: A pointwise condition for an infinitely differentiable function of several variables to be a polynomial, *Journal of Mathematical Analysis and Applications* **151**, 17–19.

H. BOHMAN [1952]: On approximation of continuous and of analytic functions, *Arkiv för Mathematik* **2**, 43–56.

H.F. BOHNENBLUST; A. SOBCZYK [1938]: Extensions of functionals on complex linear spaces. *Bulletin of the American Mathematical Society* **44**, 91–93.

O. BOLZA [1946]: *Lectures on the Calculus of Variations*, Chelsea Publishing Company, New York.

O. BONNET [1848]: Mémoire sur la théorie générale des surfaces, *Journal de l'Ecole Polytechnique* **19**, 1–146.

W.M. BOOTHBY [1975]: *An Introduction to Differentiable Manifolds and Riemannian Geometry*, Academic Press, New York.

W. BORCHERS; H. SOHR [1990]: On the equations rot $v = g$ and div$\,u = f$ with zero boundary conditions, *Hokkaido Mathematical Journal* **19**, 67–87.

K. BORSUK [1933]: Drei Sätze über die n-dimensionale euklidische Sphäre, *Fundamenta Mathematicae* **21**, 177–190.

N. BOURBAKI [1966a]: *Éléments de Mathématique. Topologie Générale; Chapitres* 1 *à* 4, Hermann, Paris (English translation: *Elements of Mathematics, General Topology: Chapters* 1–4, Springer, New York, 1998).

N. BOURBAKI [1966b]: *Eléments de Mathématique. Topologie Générale: Chapitres* 5 *à* 10, Hermann, Paris (English translation: *Elements of Mathematics, General Topology: Chapters* 5-10, Addison-Wesley, Reading, MA, 1966).

N. BOURBAKI [1970]: *Éléments de Mathématique. Théorie des Ensembles*, Hermann, Paris (English translation: *Theory of Sets*, Springer, New York, 2004).

N. BOURBAKI [1974]: *Éléments d'Histoire des Mathématiques*, Hermann, Paris (English translation: *Elements of the History of Mathematics*, Springer, New York, 1998).

J. BOURGAIN [1977]: On dentability and the Bishop-Phelps property, *Israel Journal of Mathematics* **28**, 268–271.

R.E. BRADLEY; C.E. SANDIFER [2009]: *Cauchy's Cours d'Analyse—An Annotated Translation*, Springer, Heidelberg.

A. Braides [2002]: Γ-*Convergence for Beginners*, Oxford University Press, Oxford, UK.

J.H. Bramble; S.R. Hilbert [1970]: Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation, *SIAM Journal on Numerical Analysis* **7**, 112–124.

J. Brandts; S. Korotov; M. Křížek [2011]: Generalization of the Zlámal condition for simplicial finite elements in $\mathbb{R}^d$, *Applied Mathematics* **56**, 417–424.

S.C. Brenner; R. Scott [2002]: *The Mathematical Theory of Finite Element Methods*, Springer, New York.

H. Brezis [1971]: Problèmes unilatéraux, *Journal de Mathématiques Pures et Appliquées* **9**, 1–168.

H. Brezis [1973]: *Opérateurs Maximaux Monotones*, North-Holland, Amsterdam.

H. Brezis [1983]: *Analyse Fonctionnelle. Théorie et Applications*, Masson, Paris.

H. Brezis [2011]: *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Springer, New York.

H. Brezis; M. Sibony [1971]: Equivalence de deux inéquations variationnelles, *Archive for Rational Mechanics and Analysis* **41**, 254–265.

H. Brezis; G. Stampacchia [1968]: Sur la régularité de la solution d'inéquations elliptiques, *Bulletin de la Société Mathématique de France* **96**, 153–180.

F. Brezzi [1974]: On the existence, uniqueness and approximation of saddle point problems arising from Lagrange multipliers, *Revue Française d'Automatique, Informatique, et Recherche Opérationnelle–Série Rouge* **8**, 129–151.

F. Brezzi; M. Fortin [1991]: *Mixed and Hybrid Finite Element Methods*, Springer, New York.

L.E.J. Brouwer [1911]: Beweis des Jordanschen Satzes für den n-dimensionalen Raum, *Mathematische Annalen* **71**, 314–319 and 598.

L.E.J. Brouwer [1912]: Übet Abbildungen von Mannigfaltigkeiten, *Mathematische Annalen* **71**, 97–115.

L.E.J. Brouwer [1912]: Beweis der Invarianz des n-dimensionalen Gebiets, *Mathematische Annalen* **71**, 305–315.

F.E. Browder [1963]: Nonlinear elliptic boundary value problems, *Bulletin of the American Mathematical Society* **69**, 862–874.

F.E. Browder [1965]: Existence and uniqueness theorems for solutions of nonlinear boundary value problems, in *Proceedings of Symposia in Applied Mathematics*, *Volume* XVII: *Applications of Nonlinear Partial Differential Equations in Mathematical Physics*, pp. 24–49, American Mathematical Society, Providence, RI.

B. van Brunt [2006]: *The Calculus of Variations*, Springer, New York.

L. Brutman [1997]: Lebesgue functions for polynomial interpolations–a survey, *Annals of Numerical Mathematics* **4**, 111-127.

V. BUNYAKOVSKIĬ: [1859]: Sur quelques inégalités concernant les intégrales aux différences finies, *Mémoires de l'Académie des Sciences de Saint-Peterbourg*, 7 *ème Série, Tome* 1, No. 9, 1–18.

B. BUTTAZZO; M. GIAQUINTA; S. HILDEBRANDT [1998]: *One-dimensional Variational Problems: An Introduction*, Clarendon Press, Oxford.

G. CANTOR [1899]: *Beiträge zur Begründung der transfiniten Mengenlehre*, Georg Olms Verlag (English translation: *Contributions to the Founding of Transfinite Numbers*, Dover, New York, 1955).

C. CARATHÉODORY [1907]: Über den Variabilitätsbereich der Fourier'schen Konstanten von positiven harmonischen Funktionen, *Rendiconti del Circolo Matematico di Palermo* **32**, 193–217.

C. CARATHÉODORY [1965]: *Calculus of Variations and Partial Differential Equations of the First Order*, Holden Day, San Francisco.

L. CARLESON [1966]: On convergence and growth of partial sums of Fourier series, *Acta Mathematica* **116**, 135–157.

M.P. DO CARMO [1976]: *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, NJ.

M.P. DO CARMO [1994]: *Differential Forms and Applications*, Universitext, Springer, Berlin (English translation of: *Formas Diferenciais e Aplições*, Instituto da Matematica, Pura e Aplicada, Rio de Janeiro, 1971).

N.L. CAROTHERS [2000]: *Real Analysis*, Cambridge University Press.

E. CARTAN [1927]: Sur la possibilité de plonger un espace riemannien donné dans un espace euclidien, *Annales de la Société Polonaise de Mathématiques* **6**, 1–7.

E. CARTAN [1928]: *Leçons sur la Géométrie des Espaces de Riemann*, Gauthier-Villars, Paris.

A.L. CAUCHY [1821]: *Cours d'Analyse de l'École Royale Polytechnique*, de Bure, Paris.

E. ČECH [1937]: On bicompact spaces, *Annals of Mathematics* **38**, 823–844.

E. CESÀRO [1906]: Sulle formole del Volterra, fondamentali nella teoria delle distorsioni elastiche, *Rendiconti Napoli* **12**, 311–321.

F. CHATELIN [1983]: *Spectral Approximation of Linear Operators*, Academic Press, New York.

W. CHEN; J. JOST [2002]: A Riemannian version of Korn's inequality, *Calculus of Variations* **14**, 517–530.

W.W. CHENEY [1966]: *Introduction to Approximation Theory*, McGraw-Hill, New York.

M. CHIPOT [2000]: *Elements of Nonlinear Analysis*, Birkhäuser, Basel.

M. CHIPOT [2002]: *ℓ Goes to Plus Infinity*, Birkhäuser, Basel.

M. CHIPOT [2009]: *Elliptic Equations: An Introductory Course*, Birkhäuser, Basel.

G. CHOQUET [1966]: *Topology*, Academic Press, New York.

Y. Choquet-Bruhat; C. de Witt-Morette; M. Dillard-Bleick [1982]: *Analysis, Manifolds and Physics, Second Edition*, North-Holland, Amsterdam (First Edition: 1977).

E.B. Christoffel [1869]: Über die Transformation der homogenen Differentialausdrücke zweiten Grades, *Journal für die Reine und Angewandte Mathematik* **70**, 46–70.

P.G. Ciarlet [1975]: *Lectures on the Finite Element Method*, Tata Institute of Fundamental Research, Bombay.

P.G. Ciarlet [1978]: *The Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam (reprinted in 2002 as SIAM Classics in Applied Mathematics, Volume 40, SIAM, Philadelphia).

P.G. Ciarlet [1978]: Interpolation error estimates for the reduced Hsieh-Clough-Tocher triangle, *Mathematics of Computation* **32**, 335–344.

P.G. Ciarlet [1980]: A justification of the von Kármán equations, *Archive for Rational Mechanics and Analysis* **73**, 349–389.

P.G. Ciarlet [1987]: *Introduction to Numerical Linear Algebra and Optimisation*, with the assistance of B. Miara and J.M. Thomas for the Exercises, Cambridge University Press, Cambridge, UK (translation of the original French edition, *Introduction à l'Analyse Numérique Matricielle et à l'Optimisation*, Masson, Paris, 1982, republished with a new presentation by Dunod, Paris, in 2007, and of *Exercices d'Analyse Numérique Matricielle et d'Optimisation, avec Solutions*, by P.G. Ciarlet, B. Miara, and J.M. Thomas, Masson, Paris, 1991, republished with a new presentation by Dunod, Paris, in 2001).

P.G. Ciarlet [1988]: *Mathematical Elasticity, Volume* I: *Three-Dimensional Elasticity*, North-Holland, Amsterdam.

P.G. Ciarlet [1991]: Basic error estimates for elliptic problems, in *Handbook of Numerical Analysis*, *Volume* II (P.G. Ciarlet & J.L. Lions, editors), pp. 17–351, North-Holland, Amsterdam.

P.G. Ciarlet [1997]: *Mathematical Elasticity, Volume* II: *Theory of Plates*, North-Holland, Amsterdam.

P.G. Ciarlet [2000]: *Mathematical Elasticity, Volume* III: *Theory of Shells*, North-Holland, Amsterdam.

P.G. Ciarlet [2003]: The continuity of a surface as a function of its two fundamental forms, *Journal de Mathématiques Pures et Appliquées* **82**, 253–274.

P.G. Ciarlet [2005]: *An Introduction to Differential Geometry with Applications to Elasticity*, Springer, Dordrecht.

P.G. Ciarlet; P. Ciarlet, Jr. [2005]: Another approach to linearized elasticity and a new proof of Korn's inequality, *Mathematical Models and Methods in Applied Sciences* **15**, 259–271.

P.G. Ciarlet; P. Destuynder [1979]: A justification of a nonlinear model in plate theory, *Computer Methods in Applied Mechanics and Engineering* **17/18**, 227–258.

P.G. Ciarlet; G. Geymonat [1982]: Sur les lois de comportement en élasticité non-linéaire compressible, *Comptes Rendus de l'Académie des Sciences de Paris, Série* II, **295**, 423–426.

P.G. Ciarlet; G. Geymonat; F. Krasucki [2012]: A new duality approach to elasticity, *Mathematical Models and Methods in Applied Sciences* **22**, 1150003.

P.G. Ciarlet; L. Gratie; O. Iosifescu; C. Mardare; C. Vallée [2007]: Another approach to the fundamental theorem of Riemannian geometry in $\mathbb{R}^3$, by way of rotation fields, *Journal de Mathématiques Pures et Appliquées* **87**, 237–252.

P.G. Ciarlet; L. Gratie; C. Mardare [2005]: A nonlinear Korn inequality on a surface, *Journal de Mathématiques Pures et Appliquées* **85**, 2–16.

P.G. Ciarlet; L. Gratie; C. Mardare [2008]: A new approach to the fundamental theorem of surface theory, *Archive for Rational Mechanics and Analysis* **188**, 457–473.

P.G. Ciarlet; O. Iosifescu [2009]: A new approach to the fundamental theorem of surface theory, by means of the Darboux-Vallée-Fortunée compatibility relation, *Journal de Mathématiques Pures et Appliquées* **91**, 384–401.

P.G. Ciarlet; F. Larsonneur [2002]: On the recovery of a surface with prescribed first and second fundamental forms, *Journal de Mathématiques Pures et Appliquées* **81**, 167–185.

P.G. Ciarlet; F. Laurent [2003]: Continuity of a deformation as a function of its Cauchy-Green tensor. *Archive for Rational Mechanics and Analysis* **167**, 255–269.

P.G. Ciarlet; V. Lods [1996]: On the ellipticity of linear membrane shell equations, *Journal de Mathématiques Pures et Appliquées* **75**, 107–124.

P.G. Ciarlet; C. Mardare [2003]: On rigid and infinitesimal rigid displacements in shell theory, *Journal de Mathématiques Pures et Appliquées* **83**, 1–15.

P.G. Ciarlet; C. Mardare [2003]: On rigid and infinitesimal rigid displacements in three-dimensional elasticity, *Mathematical Models and Methods in Applied Sciences* **13**, 1589–1598.

P.G. Ciarlet; C. Mardare [2004]: Continuity of a deformation in H^1 as a function of its Cauchy-Green tensor in L^1, *Journal of Nonlinear Science* **14**, 415–427.

P.G. Ciarlet; C. Mardare [2004]: Recovery of a manifold with boundary and its continuity as a function of its metric tensor, *Journal de Mathématiques Pures et Appliquées* **83**, 811–843.

P.G. Ciarlet; C. Mardare [2005]: Recovery of a surface with boundary and its continuity as a function of its two fundamental forms, *Analysis and Applications* **3**, 99–117.

P.G. Ciarlet; C. Mardare [2012]: The Newton-Kantorovich theorem, *Analysis and Applications* **10**, 249–269.

P.G. Ciarlet; S. Mardare [2001]: On Korn's inequalities in curvilinear coordinates, *Mathematical Models and Methods in Applied Sciences* **11**, 1379–1391.

P.G. Ciarlet; J. Nečas [1987]: Injectivity and self-contact in nonlinear elasticity, *Archive*

for Rational Mechanics and Analysis **97**, 171–188.

P.G. CIARLET; P. RABIER [1980]: *Les Equations de von Kármán*, Lecture Notes in Mathematics, Volume 826, Springer, Berlin.

P.G. CIARLET, P.A. RAVIART [1972]: General Lagrange and Hermite interpolation in $\mathbb{R}^n$ with applications to finite element methods, *Archive for Rational Mechanics and Analysis* **46**, 177–199.

P.G. CIARLET; E. SANCHEZ-PALENCIA [1996]: An existence and uniqueness theorem for the two-dimensional linear membrane shell equations, *Journal de Mathématiques Pures et Appliquées* **75**, 51–67.

P.G. CIARLET; M.H. SCHULTZ; R.S. VARGA [1969]: Numerical methods of high-order accuracy for nonlinear boundary value problems V: Monotone operator theory, *Numerische Mathematik* **13**, 51–79.

P.G. CIARLET; C. WAGSCHAL [1971]: Multipoint Taylor formulas and applications to the finite element method, *Numerische Mathematik* **17**, 84–100.

D. CIORANESCU; P. DONATO [1999]: *An Introduction to Homogenization*, Oxford Lecture Series in Mathematics and Its Applications, Volume 17, Oxford University Press, Oxford, UK.

D. CIORANESCU; P. DONATO; M.P. ROQUE [2012]: *Introduction to Classical and Variational Partial Differential Equations*, The University of the Philippines Press, Quezon City.

D. CIORANESCU; J. SAINT JEAN PAULIN [1999]: *Homogenization of Reticulated Structures*, Applied Mathematical Sciences, Volume 136, Springer, Berlin.

J.A. CLARKSON [1936]: Uniformly convex spaces, *Transactions of the American Mathematical Society* **40**, 396–414.

C. COATMÉLEC [1966]: Approximation et interpolation des fonctions différentiables de plusieurs variables, *Annales Scientifiques de l'Ecole Normale Supérieure* **83**, 271–341.

D. CODAZZI [1868–1869]: Sulle coordinate curvilinee d'una superficie dello spazio, *Annali di Mathematica Pura e Applicata* **2**, 101–119.

E.A. CODDINGTON; N. LEVINSON [1955]: *Theory of Ordinary Differential Equations*, McGraw Hill, New York.

P.J. COHEN [1963]: The independence of the continuum hypothesis, *Proceedings of the National Academy of Sciences, USA* **50**, 1143–1148.

P.J. COHEN [1964]: The independence of the continuum hypothesis, *Proceedings of the National Academy of Sciences, USA* **51**, 105–110.

P.J. COHEN [1966]: *Set Theory and the Continuum Hypothesis*, Benjamin, New York.

B.D. COLEMAN; W. NOLL [1959]: On the thermostatics of continuous media, *Archive for Rational Mechanics and Analysis* **4**, 97–128.

P. CONSTANTIN; C. FOIAS [1988]: *Navier-Stokes Equations*, University of Chicago Press, Chicago, IL.

J. Conway [1990]: *A Course in Functional Analysis, Second Edition*, Springer, New York (First Edition: 1985).

E. Corominas; F.S. Balaguer [1954]: Conditions for an infinitely differentiable function to be a polynomial (title in Spanish), *Revista Matemática Hispano-Americana* **14**, 26–43.

E. Cosserat; F. Cosserat [1896]: Sur la théorie de l'élasticité. Premier mémoire, *Annales de la Faculté des Sciences de l'Université de Toulouse* **10**, 1–116.

R. Courant [1920]: Über die Eigenwerte bei den Differentialgleichungen der Mathematischen Physik, *Mathematische Zeitschrift* **7**, 1–57.

M. Crouzeix; A.L. Mignot [1983]: *Analyse Numérique des Equations Différentielles*, Masson, Paris.

G. Csato; B. Dacorogna; O. Kneuss [2011]: *The Pullback Equation*, Birkhäuser, Basel.

H. Curien; M. Schmidt, editors [1990]: *Hommes de Science*, Hermann, Paris.

B. Dacorogna [1982]: Minimal hypersurfaces in parametric form with nonconvex integrands, *Indiana University Mathematics Journal* **31**, 531–552.

B. Dacorogna [2010]: *Direct Methods in the Calculus of Variations, Second Edition*, Springer, Berlin (First Edition: 1989).

G. Dal Maso [1993]: *An Introduction to Γ-Convergence*, Birkhäuser, Boston.

R. Dautray; J.L. Lions [2000a]: *Mathematical Analysis and Numerical Methods for Science and Technology, Volume* 1: *Physical Origins and Classical Methods*, Springer, Heidelberg.[103)]

R. Dautray; J.L. Lions [2000b]: *Mathematical Analysis and Numerical Methods for Science and Technology, Volume* 2: *Functional and Variational Methods*, Springer, Heidelberg.

R. Dautray; J.L. Lions [2000c]: *Mathematical Analysis and Numerical Methods for Science and Technology, Volume* 3: *Spectral Theory and Applications*, Springer, Heidelberg.

R. Dautray; J.L. Lions [2000d]: *Mathematical Analysis and Numerical Methods for Science and Technology, Volume* 4: *Integral Equations and Numerical Methods*, Springer, Heidelberg.

R. Dautray; J.L. Lions [2000e]: *Mathematical Analysis and Numerical Methods for Science and Technology, Volume* 5: *Evolution Problems* I, Springer, Heidelberg.

R. Dautray; J.L. Lions [2000f]: *Mathematical Analysis and Numerical Methods for Science and Technology, Volume* 6: *Evolution Problems* II, Springer, Heidelberg.

C. Davis [1971]: The Toeplitz-Hausdorff theorem explained, *Canadian Mathematical Bulletin* **14**, 245–246.

P.J. Davis [1963]: *Interpolation and Approximation*, Dover, New York.

P.J. Davis; P. Rabinowitz [1975]: *Methods of Numerical Integration*, Academic Press, New York.

[103)] These six volumes are translated from *Analyse Mathématique et Calcul Numérique pour les Sciences et les Techniques*, Masson, Paris et Commissariat à l'Energie Atomique, Paris, 1984—1985.

L. DEBNATH; P. MIKUSIŃSKI [1999]: *Hilbert Spaces with Applications, Second Edition*, Academic Press, New York (First Edition: 1990).

J.P. DEDIEU [2006]: *Points Fixes, Zéros et la Méthode de Newton*, Springer, Berlin.

E. DE GIORGI [1975]: Sulla convergenza di alcune successioni di integrali del tipo dell'area, *Rendiconti Mathematica Roma* **8**, 227–294.

E. DE GIORGI [1977]: Γ-convergenza e G-convergenza, *Bolletina Unione Mathematica Italiana* **5**, 213–220.

E. DE GIORGI; G. DAL MASO [1983]: Γ-*Convergence and Calculus of Variations*, Lecture Notes in Mathematics, Volume 979, Springer, Berlin.

K. DEIMLING [1985]: *Nonlinear Functional Analysis*, Springer, Berlin.

L. DEMKOWICZ [2000]: Babuška ⇔ Brezzi??, Technical Report, Texas Institute for Computational and Applied Mathematics, TICAM Seminar (October 31, 2000).

Z. DENKOWSKI; S. MIGÓRSKI; N.S. PAPAGEORGIOU [2003]: *An Introduction to Nonlinear Analysis: Applications*, Kluwer, Boston.

P. DEUFLHARD [2004]: *Newton Methods for Nonlinear Problems—Affine Invariance and Adaptive Algorithms*, Springer, Berlin.

R. DEVORE, G.G. LORENTZ [1993]: *Constructive Approximation*, Springer, Heidelberg.

E. DIBENEDETTO [2002]: *Real Analysis*, Birkhäuser, Boston.

E. DIBENEDETTO [2010]: *Partial Differential Equations, Second Edition*, Birkhäuser, Boston (First Edition: 1995, Springer, New York).

J. DIESTEL [1975]: *Geometry of Banach Spaces: Selected Topics*, Springer, Berlin.

J. DIEUDONNÉ [1950]: Deux exemples singuliers d'équations différentielles, *Acta Scientiarum Mathematicarum B (Szeged)* **12**, 38–40.

J. DIEUDONNÉ [1960]: *Foundations of Modern Analysis*, Academic Press, New York.

J. DIEUDONNÉ [1981]: *History of Functional Analysis*, North-Holland, Amsterdam.

J. DIEUDONNÉ [1989]: *A History of Algebraic and Differential Topology*, 1900–1960, Birkhäuser, Boston.

G. DINCA [2001]: A Fredholm-type result for a couple of nonlinear operators, *Comptes Rendus de l'Académie des Sciences de Paris, Série* 1, **333**, 4015–4019.

G. DINCA [2004]: Duality mappings on infinite dimensional reflexive and smooth Banach spaces are not compact, *Bulletin de l'Académie Royale de Belgique, Classes des Sciences* **6**, 33–40.

G. DINCA; P. JEBELEAN; J. MAWHIN [2001]: Variational and topological methods for Dirichlet problems with p-Laplacian, *Portugaliae Mathematica* **58**, 339–378.

G. DINCA; J. MAWHIN [2013]: *Brouwer Degree and Applications*, to appear.

U. DINI [1878]: *Analisi Infinitesimale. Lezioni dettate nella Reale Università di Pisa, Anno Accademico* 1877–1878.

U. DINI [1878]: *Fondamenti per la Teoria delle Funzioni di Variabili Reali*, T. Nistri, Pisa.

J. DIXMIER [1953]: Sur les bases orthonormales dans les espaces préhilbertiens, *Acta Scientiarum Mathematicarum Szeged* **15**, 29–30.

L. DONATI [1890]: Illustrazione al teorema del Menabrea, *Memorie della Accademia delle Scienze dell'Istituto di Bologna* **10**, 267–274.

L. DONATI [1894]: Ulteriori osservazioni intorno al teorema del Menabrea, *Memorie della Accademia delle Scienze dell'Istituto di Bologna* **4**, 449–474.

P. DU BOIS-RAYMOND [1876]: Untersuchungen über die Convergenz und Divergenz der Fourierschen Darstellungsformeln, *Abhandlungen der Mathematisch-Physikalischen Klasse der Königlich Bayerischen Akademie der Wissenschaften* **12**, 1–103.

R.M. DUDLEY [1964]: On sequential convergence, *Transactions of the American Mathematical Society* **112**, 483–507.

J. DUISTERMAAT; J.A. KOLK [2010]: *Distributions: Theory and Applications*, Springer, New York.

N. DUNFORD; J. SCHWARTZ [1958]: *Linear Operators, Part* I: *General Theory*, Interscience, New York (Reprinting: Wiley Classics Library, 1988).

N. DUNFORD; J. SCHWARTZ [1963]: *Linear Operators, Part* II: *Spectral Theory—Self Adjoint Operators in Hilbert Spaces*, Interscience, New York (Reprinting: Wiley Classics Library, 1988).

N. DUNFORD; J. SCHWARTZ [1971]: *Linear Operators, Part* III: *Spectral Operators*, Interscience, New York (Reprinting: Wiley Classics Library, 1988).

G. DUVAUT; J.L. LIONS [1976]: *Inequalities in Mechanics and Physics*, Springer, Berlin (translation of the original French edition, *Les Inéquations en Mécanique et en Physique*, Dunod, Paris, 1972).

W.F. EBERLEIN [1947]: Weak compactness in Banach spaces I, *Proceedings of the National Academy of Sciences, USA* **33**, 51–53.

A. EISINBERG; G. FEDELE; G. FRANZÈ [2004]: Lebesgue constant for Lagrange interpolation on equidistant nodes, *Analysis in Theory and Applications* **20**, 323–331.

I. EKELAND [1974]: On the variational principle, *Journal of Mathematical Analysis and Applications* **47**, 324–353.

I. EKELAND [1979]: Nonconvex minimization problems, *Bulletin of the American Mathematical Society* **1**, 443–473.

I. EKELAND; R. TÉMAM [1976]: *Convex Analysis and Variational Problems*, North-Holland, Amsterdam (reprinted in 1999 as SIAM Classics in Applied Mathematics, Volume 28; translation of the original French edition, *Analyse Convexe et Problèmes Variationnels*, Dunod, Paris, 1974).

L. EULER [1775]: On representations of a spherical surface on the plane, *Proceedings of the Saint Petersburg Academy of Sciences*.

L.C. EVANS [2010]: *Partial Differential Equations, Second Edition*, American Mathematical Society, Providence, RI (First Edition: 1998).

L.C. EVANS; R.F. GARIEPY [1992]: *Measure Theory and Fine Properties of Functions*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL.

Ky FAN [1953]: Minimax theorems, *Proceedings of the National Academy of Sciences* **39**, 42–47.

J. FARKAŠ [1901]: Theorie der einfachen Ungleichungen, *Journal für die Reine und Angewandte Mathematik* **124**, 1–27.

H. FEDERER [1969]: *Geometric Measure Theory*, Springer, New York.

L. FEJÉR [1900]: Sur les fonctions bornées et intégrables, *Comptes Rendus de l'Académie des Sciences, Paris* **131**, 984–987.

W. FENCHEL [1949]: On conjugate convex functions, *Canadian Journal of Mathematics* **1**, 73–77.

G. FICHERA [1964]: Problemi elastostatici con vincoli unilaterali: il problema de Signorini con ambigue condizioni al contorno, *Memorie dell'Accademia Nazionale dei Lincei* **8**, 91–140.

G. FICHERA [1972a]: Existence theorems in elasticity, in *Handbuch der Physik* VIa/2 (S. FLÜGGE & C. TRUESDELL, editors), pp. 347–389, Springer, Berlin.

G. FICHERA [1972b]: Boundary value problems of elasticity with unilateral constraints, in *Handbuch der Physik* VIa/2 (S. FLÜGGE & C. TRUESDELL, editors), pp. 391–424, Springer, Berlin.

E. FISCHER [1905]: Über quadratische Formen mit reellen Koeffizienten, *Monatshefte für Mathematik und Physik* **16**, 234–249.

E. FISCHER [1907]: Sur la convergence en moyenne, *Comptes Rendus de l'Académie des Sciences* **144**, 1022–1024.

S.R. FOGUEL [1958]: On a theorem of A.E. Taylor, *Proceedings of the American Mathematical Society* **9**, 325.

C. FOIAS; O. MANLEY; R. ROSA; R. TEMAM [2001]: *Navier-Stokes Equations and Turbulence*, Cambridge University Press, Cambridge, UK.

G.B. FOLLAND [1984]: *Real Analysis*, Wiley, New York.

I. FONSECA; W. GANGBO [1995]: *Degree Theory in Analysis and Applications*, Clarendon Press, Oxford, UK.

L.E. FRAENKEL [2000]: *An Introduction to Maximum Principle and Symmetry in Elliptic Problems*, Cambridge University Press, Cambridge, UK.

S.P. FRANKLIN [1965]: Spaces in which sequences suffice, *Fundamenta Mathematicae* **57**, 107–115.

S.P. FRANKLIN [1967]: Spaces in which sequences suffice, *Fundamenta Mathematicae* **61**, 51–56.

T.G. FREEMAN [2002]: *Portraits of the Earth. A Mathematician Looks at Maps*, American Mathematical Society, Providence.

K.O. FRIEDRICHS [1947]: On the boundary-value problems of the theory of elasticity and Korn's inequality, *Annals of Mathematics* **48**, 441–471.

K.O. FRIEDRICHS [1981]: *Spectral Theory of Operators in Hilbert Spaces*, Springer, Berlin.

G. FRIESECKE; R.D. JAMES; M.G. MORA; S. MÜLLER [2003]: Derivation of nonlinear bending theory for shells from three-dimensional nonlinear elasticity by Gamma-convergence, *Comptes Rendus de l'Académie des Sciences de Paris, Série* 1, **336**, 697–702.

G. FRIESECKE; R.D. JAMES; S. MÜLLER [2002]: A theorem on geometric rigidity and the derivation of nonlinear plate theory from three dimensional elasticity, *Communications on Pure and Applied Mathematics* **55**, 1461–1506.

G. FRIESECKE; R.D. JAMES; S. MÜLLER [2006]: A hierarchy of plate models derived from nonlinear elasticity by Gamma-convergence, *Archive for Rational Mechanics and Analysis* **180**, 183–236.

G. FROBENIUS [1912]: Über Matrizen aus nicht negativen Elementen, *Sitzungsberichte Preußische Akademie der Wissenschaft*, Berlin, 456–477.

B. GALERKIN [1915]: *Rods and Plates*, Vestnik Inženerov **19** (in Russian).

S. GALLOT; D. HULIN; J. LAFONTAINE [2004]: *Riemannian Geometry, Third Edition*, Springer, Berlin (First Edition: 1987).

F.R. GANTMACHER [1959]: *The Theory of Matrices, Volumes* 1 *and* 2, Chelsea, New York.

R. GÂTEAUX [1919]: Fonctions d'une infinité de variables indépendantes, *Bulletin de la Société Mathématique de France* **47**, 70–96.

C.F. GAUSS [1809]: *Theoria Motus Corporum Coelestium in Sectionibus Conicis Solum Ambientium*, Perthes und Besser, Hamburg.

C.F. GAUSS [1822]: Anwendung der Wahrscheinlichkeitsrechnung auf eine Aufgabe der practischen Geometrie, *Astronomische Nachrichten* **1**, 81–86.

C.F. GAUSS [1827]: Disquisitiones generales circa superficies curvas, *Commentationes Societatis Regiae Scientiarum Gottingensis Recentiores* **6**, 99–146.

C.F. GAUSS [1828]: Disquisitiones generales circas superficies curvas, *Commentationes societatis regiae scientiarum Gottingensis recentiores* **6**, Göttingen.

G. GEYMONAT; F. KRASUCKI [2005]: Some remarks on the compatibility conditions in elasticity, *Accademia Nazionale delle Scienze detta dei* XL. *Rendiconti. Serie* V. *Memorie di Matematica e Applicazioni. Parte* I, **29**, 175–181.

G. GEYMONAT; F. KRASUCKI [2006]: Beltrami's solutions of general equilibrium equations in continuum mechanics, *Comptes Rendus de l'Académie des Sciences de Paris, Série* 1, **342**, 359–363.

G. GEYMONAT; G. GILARDI [1998]: Contre-exemple à l'inégalité de Korn et au lemme de Lions dans des domaines irréguliers, in *Equations aux Dérivées Partielles et Applications.*

Articles Dédiés à Jacques-Louis Lions, pp. 541–548, Gauthier-Villars, Paris.

G. GEYMONAT; P. SUQUET [1986]: Functional spaces for Norton-Hoff materials, *Mathematical Methods in the Applied Sciences* **8**, 206–222.

M. GHERGU; V.D. RĂDULESCU [2008]: *Singular Elliptic Problems: Bifurcation and Asymptotic Analysis*, Clarendon Press, Oxford, UK.

M. GHERGU; V.D. RĂDULESCU [2012]: *Nonlinear PDEs—Mathematical Models in Biology, Chemistry and Population Genetics*, Springer, Heidelberg.

M. GIAQUINTA; S. HILDEBRANDT [2006a]: *Calculus of Variations* I: *The Lagrangian Formalism*, Springer, New York.

M. GIAQUINTA; S. HILDEBRANDT [2006b]: *Calculus of Variations* II: *The Hamiltonian Formalism*, Springer, New York.

D. GILBARG; N.S. TRUDINGER [1998]: *Elliptic Partial Differential Equations, Revised Second Edition*, Springer, Berlin (First Edition: 1977).

V. GIRAULT; P.A. RAVIART [1979]: *Finite Element Approximation of the Navier-Stokes Equations*, Lecture Notes in Mathematics, Volume 749, Springer, Berlin.

V. GIRAULT; P.A. RAVIART [1986]: *Finite Element Methods for Navier-Stokes Equations*, Springer, Berlin.

E. GIUSTI [1984]: *Minimal Surfaces and Functions of Bounded Variations*, Birkhäuser, Boston.

E. GIUSTI [2003]: *Direct Methods in the Calculus of Variations*, World Scientific, Singapore.

R. GLOWINSKI [1984]: *Numerical Methods for Nonlinear Variational Problems*, Springer, New York.

R. GLOWINSKI [2003]: Finite element methods for incompressible viscous flows, in *Handbook of Numerical Analysis, Volume* IX (P.G. CIARLET & J.L. LIONS, editors), pp. 3–1176, North-Holland, Amsterdam.

R. GLOWINSKI; H. LANCHON [1973]: Torsion élasto-plastique d'une barre cylindrique de section multi-connexe, *Journal de Mécanique* **12**, 151–171.

R. GLOWINSKI; J.L. LIONS; R. TRÉMOLIÈRES [1981]: *Numerical Analysis of Variational Inequalities*, North-Holland, Amsterdam (translation of the original French edition, *Analyse Numérique des Inéquations Variationnelles*, Dunod, Paris, 1976).

J. GOBERT [1962]: Une inégalité fondamentale de la théorie de l'élasticité, *Bulletin de la Société Royale des Sciences de Liège* **31**, 182–191.

K. GÖDEL [1940]: *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory*, Princeton University Press, Princeton, NJ.

C. GOFFMAN; G. PEDRICK [1965]: *First Course in Functional Analysis*, Prentice-Hall, Englewood Cliffs, NJ.

H.H. GOLDSTINE [1980]: *A History of the Calculus of Variations from the* 17*th to the* 19*th Century*, Springer, New York.

W.B. GRAGG; R.A. TAPIA [1974]: Optimal error bounds for the Newton-Kantorovich theorem, *SIAM Journal on Numerical Analysis* **11**, 10–13.

J.P. GRAM [1883]: Über die Entwicklung reeller Funktionen in Reihen mittelst der Methode der kleinsten Quadrate, *Journal für die Reine und Angewandte Mathematik* **94**, 41–73.

J. GRAY [2012]: *Henry Poincaré: A Scientific Biography*, Princeton University Press, Princeton, NJ.

P. GRISVARD [1992]: *Singularities in Boundary Value Problems*, Masson, Paris.

W. GROMES [1981]: Ein einfacher Beweis des Satzes von Borsuk, *Mathematische Zeitschrift* **178**, 399–400.

M.E. GURTIN [1972]: The linear theory of elasticity, in *Handbuch der Physik*, *Volume* VIa/2 ((S. FLÜGGE & C. TRUESDELL, editors), pp. 1–295, Springer, Berlin.

M.E. GURTIN [1981]: *Topics in Finite Elasticity*, CBMS-NSF Regional Conference Series in Applied Mathematics, Volume 35, SIAM, Philadelphia.

Dzung Minh HA [2007]: *Functional Analysis: A Gentle Introduction*, Matrix Editions, Ithaca, NY.

A. HAAR [1918]: Die Minkowskische Geometrie und die Annäherung an stetige Funktionen, *Mathematische Annalen* **78**, 294–311.

J. HADAMARD [1902]: Sur les problèmes aux dérivées partielles et leur signification physique, *Princeton University Bulletin* **13**, 49–52.

H. HAHN [1927]: Über lineare Gleichungssysteme in linearen Räumen, *Journal de Crelle* **157**, 214–229.

P.R. HALMOS [1950]: *Measure Theory*, Van Nostrand, Princeton, NJ.

P.R. HALMOS [1970]: How to write mathematics, *L'Enseignement Mathématique* **16**, 123–152.

P.R. HALMOS [1974]: *A Hilbert Space Problem Book, Second Edition*, Springer, New York (First Edition: 1960).

P.R. HALMOS [1985]: *I Want to Be a Mathematician*, Springer, New York.

P.R. HALMOS [1987]: *I Have a Photographic Memory*, American Mathematical Society, Providence, RI.

G. HAMEL [1905]: Eine Basis aller Zahlen und die unstetigen Lösungen der Funktionalgleichung $f(x+y)=f(x)+f(y)$, *Mathematische Annalen* **60**, 459–462.

G.H. HARDY [1916]: Weierstrass's non-differentiable function, *Transactions, American Mathematical Society* **17**, 301–325.

G.H. HARDY [1925]: Notes on some points in the integral calculus. LX. An inequality between integrals, *Messengers of Mathematics* **54**, 150–156.

P. HARTMAN [2002]: *Ordinary Differential Equations, Second Edition*, SIAM, Philadelphia (First Edition: 1964, John Wiley & Sons, New York).

P. HARTMAN; G. STAMPACCHIA [1966]: On some nonlinear elliptic differential functional

equations, *Acta Mathematica* **115**, 271–310.

P. HARTMAN; A. WINTNER [1950]: On the embedding problem in differential geometry, *American Journal of Mathematics* **72**, 553–564.

P. HARTMAN; A. WINTNER [1950]: On the fundamental equations of differential geometry, *American Journal of Mathematics* **72**, 757–774.

F. HAUSDORFF [1919]: Der Wertvorrat einer Bilinearform, *Mathematische Zeitschrift* **3**, 314–316.

E. HEINZ [1959]: An elementary analytic theory of the degree of mapping in n-dimensional space, *Journal of Mathematics and Mechanics* **8**, 231–247.

E. HELLINGER; O. TOEPLITZ [1910]: Grundlagen für eine Theorie der unendlichen Matrizen, *Mathematische Annalen* **69**, 281–330.

M. HÉNON [1976]: A two-dimensional mapping with a strange attractor, *Communications in Mathematics and Physics* **50**, 69–77.

C. HERMITE [1878]: Sur la formule d'interpolation de Lagrange, *Journal für die reine und angewandte Mathematik* **84**, 70–79.

M.R. HESTENES [1975]: *Optimization Theory—The Finite Dimensional Case*, John Wiley, New York.

E. HEWITT; K. STROMBERG [1965]: *Real and Abstract Analysis—A Modern Treatment of the Theory of Functions of a Real Variable*, Springer, New York.

E. HILLE; C. SZEGÖ; J.D. TAMARKIN [1937]: On some generalizations of a theorem of A. Markoff, *Duke Mathematical Journal* **8**, 729–739.

J.B. HIRIART-URRUTY; C. LEMARÉCHAL [1993a]: *Convex Analysis and Minimization Algorithms* I: *Fundamentals*, Springer, Berlin.

J.B. HIRIART-URRUTY; C. LEMARÉCHAL [1993b]: *Convex Analysis and Minimization Algorithms* II: *Advanced Theory and Bundle Methods*, Springer, Berlin.

O. HÖLDER [1889]: *Über einen Mittelwertsatz*, Göttinger Nachrichten, 38–47.

E. HOPF [1927]: Elementare Bemerkungen über die Lösungen partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus, in *Sitzungsberichte der Preußischen Akademie der Wissenschaften, Berlin*, 147–152.

C.O. HORGAN [1995]: Korn's inequalities and their applications in continuum mechanics, SIAM Review **37**, 491–511.

L. HÖRMANDER [1955]: On the theory of general partial differential operators, *Acta Mathematica* **94**, 161–248.

L. HÖRMANDER [1983]: *The Analysis of Partial Differential Operators, Volume* 1, Springer, New York.

R.A. HORN; C.R. JOHNSON [1985]: *Matrix Analysis*, Cambridge University Press, Cambridge, UK.

R.A. HORN; C.R. JOHNSON [1991]: *Topics in Matrix Analysis*, Cambridge University Press,

Cambridge, UK.

A.S. Householder [1964]: *The Theory of Matrices in Numerical Analysis*, Blaisdell, New York.

J.L.W.V. Jensen [1906]: Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Mathematica* **30**, 175–193.

H. Jacobowith [1982]: *The Gauss-Codazzi equations, Tensor* (*N.S.*) **39**, 15–22.

E. Jakimowicz; A. Miranovicz, editors [2011]: *Stefan Banach: Remarkable Life, Brilliant Mathematics*, American Mathematical Society, Providence, RI.

R.C. James [1951]: A non-reflexive Banach space isometric with its second conjugate space, *Proceedings of the National Academy of Sciences, USA* **37**, 174–177.

R.C. James [1964]: Characterizations of reflexivity, *Studia Mathematica* **23**, 205–216.

R.C. James [1972]: Reflexivity and the sup of linear functionals, *Israel Journal of Mathematics* **13**, 289–301.

P. Jamet [1976]: Estimation d'erreur pour des éléments finis droits presque dégénérés, *Revue Française d'Automatique, Informatique, Recherche Opérationnelle, Série Rouge: Analyse Numérique* **10**, 43–61.

M. Janet [1926]: Sur la possibilité de plonger un espace riemannien donné dans un espace euclidien, *Annales de la Société Polonaise de Mathématiques* **5**, 38–43.

D. Jerison; C.E. Kenig [1995]: The inhomogeneous Dirichlet problem in Lipschitz domains, *Journal of Functional Analysis* **130**, 161–219.

J. Jost [2005]: *Postmodern Analysis*, Springer, Berlin.

Y. Kannai [1981]: An elementary proof of the no-retraction theorem, *American Mathematical Monthly* **88**, 264–268.

L.V. Kantorovich [1948]: Functional analysis and applied mathematics, *Uspehi Matematiceskii Nauk* (*New Series*) **3**, 89–185 (in Russian).

L.V. Kantorovich; G.P. Akilov [1959]: *Functional Analysis in Normed Vector Spaces*, Fizmatgiz, Moscow (in Russian) (English translation: Pergamon, New York, 1964).

S. Karlin [1959]: Positive operators, *Journal of Mathematics and Mechanics* **8**, 907–937.

T. Von Kármán [1910]: Festigkeitsprobleme im Maschinenbau, in *Encyclopädie der Mathematischen Wissenschaften, Volume* IV/4, pp. 311–385, Leipzig.

T. Kato [1966]: *Perturbation Theory for Linear Operators*, Springer, Berlin (Corrected Printing of the Second Edition: 1980).

O. Kavian [1993]: *Introduction à la Théorie des Points Critiques et Applications aux Problèmes Elliptiques*, Springer, Paris.

J. Kelley [1955]: *General Topology*, Van Nostrand, Princeton, NJ.

O.D. Kellogg [1929]: *Foundations of Potential Theory*, Springer, Berlin.

S. Kesavan [1989]: *Topics in Functional Analysis and Applications*, Wiley, New Delhi.

S. KESAVAN [2004]: *Nonlinear Functional Analysis—A First Course*, Hindustan Book Agency, Gurgaon.

S. KESAVAN [2005]: On Poincaré's and J.L. Lions' lemmas, *Comptes Rendus de l'Académie des Sciences de Paris, Série* I, **340**, 27–30.

S. KESAVAN [2006]: *Symmetrization & Applications*, World Scientific, Singapore.

S. KESAVAN [2009]: *Functional Analysis*, Hindustan Book Agency, Gurgaon.

D. KINDERLEHRER; G. STAMPACCHIA [1980]: *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York (reprinted as Classics in Applied Mathematics, Volume 31, SIAM, Philadelphia, 2002).

J. KISYNSKI [1959]: Convergence du type *L, Colloquium Mathematicum* **7**, 205–211.

W. KLINGENBERG [1973]: *Eine Vorlesung über Differentialgeometrie*, Springer, Berlin (English translation: *A Course in Differential Geometry*, Springer, Berlin, 1978).

A.W. KNAPP [2005a]: *Basic Real Analysis*, Birkhäuser, Boston.

A.W. KNAPP [2005b]: *Advanced Real Analysis*, Birkhäuser, Boston.

V.I. KONDRACHOV [1945]: Certain properties of functions in the spaces L^p, *Doklady Akademii Nauk SSSR* **48**, 535–538.

V.A. KONDRAT'EV; O.A. OLEINIK [1988]: Boundary-value problems for the system of elasticity theory in unbounded domains. Korn's inequalities, *Uspehi Mathematiceskii Nauk* **43**, 55–98 (in Russian) [English translation: *Russian Mathematical Surveys* **43** (1988), 65–119].

A. KORN [1906]: Die Eigenschwingungen eines elastischen Körpers mit ruhender Oberfläche, *Sitzungsberichte der Mathematisch-physikalischen Klasse der Königlich bayerischen Akademie der Wissenschaften zu München* **36**, 351–402.

A. KORN [1908]: Solution générale du problème d'équilibre dans la théorie de l'élasticité, dans le cas où les efforts sont donnés à la surface, *Annales de la Faculté des Sciences de Toulouse* **10**, 165–269.

A. KORN [1909]: Über einige Ungleichungen, welche in der Theorie der elastischen und elektrischen Schwingungen eine Rolle spielen, *Bulletin International de l'Académie des Sciences de Cracovie* **9**, 705–724.

P.P. KOROVKIN [1959]: *Linear Operators and Approximation Theory*, Fitzmatgiz, Moscow (in Russian) [English translation, Hindustan Publishing Corporation, Delhi, 1960].

P.P. KOROVKIN [1959]: On convergence of linear positive operators in the space of continuous functions, *Doklady Akademii Nauk SSR* **90**, 961–964 (in Russian).

I. KRA; S.R. SIMANCA [2012]: On circulant matrices, *Notices of the American Mathematical Society* **59**, 368–377.

S.G. KRANTZ [2004]: *Real Analysis and Foundations, Second Edition*, Studies in Advanced Mathematics, Chapman & Hall/CRC, Boca Raton, FL (First Edition: 1991).

M.A. KRASNOSELSKII [1960]: Fixed points of cone-compressive or cone-extending operators,

Soviet Mathematics Doklady **1**, 1285–1288.

M. KREIN; M. RUTMAN [1948]: Linear operators leaving invariant a cone in a Banach space, *Uspehi Mathematiceskii Nauk* **3**, 3–95 [in Russian; English translation: *American Mathematical Society Translations* 1950, No. 26].

E. KREYSZIG [1978]: *Introductory Functional Analysis with Applications*, John Wiley, New York (reprinted in the *Wiley Classics Library Edition*, 1989).

B. KRIPKE [1967]: One more reason why sequences are not enough, *American Mathematical Monthly* **74**, 563–565.

A. KUFNER; L. MALIGRANDA; L.E. PERSSON [2007]: *The Hardy Inequality: About Its History and Some Related Results*, Vydavatelský Servis, Pilsen.

H.W. KUHN; A.W. TUCKER [1951]: Nonlinear programming, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (J. NEYMAN, editor), pp. 481–492, University of California Press, Berkeley.

W. KÜHNEL [2002]: *Differentialgeometrie*, Fried. Vieweg & Sohn, Wiesbaden (English translation: *Differential Geometry: Curves-Surfaces-Manifolds*, American Mathematical Society, Providence, RI, 2002).

O.A. LADYZHENSKAYA [1969]: *The Mathematical Theory of Viscous Flows, Second Edition*, Gordon and Breach, New York.

J.L. LAGRANGE [1760]: Essai d'une nouvelle méthode pour déterminer les maxima et les minima des formules intégrales indéfinies, *Miscellanea Taurinensia* **325**, 173–199.

J.L. LAGRANGE [1773]: Solutions analytiques de quelques problèmes sur les pyramides triangulaires, *Mémoire de l'Académie Royale de Berlin.*

J.L. LAGRANGE [1812]: Leçons élémentaires de mathématiques données à l'Ecole Normale en 1795, *Journal de l'Ecole Polytechnique*, VIIe et VIIIe cahiers, t-II.

S. LANG [1993]: *Real and Functional Analysis, Third Edition*, Springer, New York.

P.S. LAPLACE [1820]: *Théorie Analytique des Probabilités, Troisième Edition, Premier Supplément: Sur l'Application du Calcul des Probabilités à la Philosophie Naturelle*, Courcier, Paris.

M. LAVRENTIEV [1926]: Sur quelques problèmes du calcul des variations, *Annales de Mathématiques Pures et Appliquées* **4**, 7–18.

D.F. LAWDEN [1989]: *Elliptic Functions and Applications*, Applied Mathematical Sciences Series, Volume 98, Springer, Heidelberg.

P.D. LAX [2002]: *Functional Analysis*, Wiley-Interscience, New York.

P.D. LAX; A.M. MILGRAM [1954]: Parabolic equations, in *Contributions to the Theory of Partial Differential Equations, Annals of Mathematics Studies, No.* 33, pp. 167–190, Princeton University Press Princeton, NJ.

L.P. LEBEDEV; M.J. CLOUD [2003]: *Tensor Analysis*, World Scientific, Singapore.

H. LEBESGUE [1901]: Sur une généralisation de l'intégrale définie, *Comptes Rendus des*

Séances de l'Académie des Sciences **132**, 1025–1027.

H. LEBESGUE [1909]: Sur les intégrales singulières, *Annales de la Faculté des Sciences de l'Université de Toulouse* **1**, 25–117.

H. LE DRET; A. RAOULT [1995]: The nonlinear membrane model as variational limit of nonlinear three-dimensional elasticity, *Journal de Mathématiques Pures et Appliquées* **74**, 549–578.

H. LE DRET; A. RAOULT [1996]: The membrane shell model in nonlinear elasticity: A variational asymptotic derivation, *Journal of Nonlinear Science* **6**, 59–94.

A.M. LEGENDRE [1805]: *Nouvelle Méthode pour la Détermination des Orbites des Comètes*, Chez Didot, Paris.

J. LERAY [1933]: Essai sur le mouvement plan d'un liquide visqueux que limitent des parois, *Journal de Mathématiques Pures et Appliquées* **13**, 331–418.

J. LERAY [1933]: Sur le mouvement d'un liquide visqueux emplissant l'espace, *Acta Mathematica* **63**, 193–248.

J. LERAY [1935]: Topologie des espaces abstraits de M. Banach, *Comptes Rendus de l'Académie des Sciences de Paris* **200**, 1082–1084.

J. LERAY [1950]: La théorie des points fixes et ses applications en analyse, in *Proceedings–International Congress of Mathematicians, Volume* 2, pp. 202–208, Cambridge.

J. LERAY; J.L. LIONS [1965]: Quelques résultats de Visik sur les problèmes elliptiques non linéaires par les méthodes de Minty-Browder, *Bulletin de la Société Mathématique de France* **93**, 97–107.

J. LERAY; J. SCHAUDER [1934]: Topologie et équations fonctionnelles, *Annales Scientifiques de l'Ecole Normale Supérieure* **51**, 45–78.

T. LEWIŃSKI; J. TELEGA [2000]: *Plates, Laminates and Shells—Asymptotic Analysis and Homogenization*, World Scientific, Singapore.

H. LEWY; G. STAMPACCHIA [1969]: On the regularity of the solution of a variational inequality, *Communications on Pure and Applied Mathematics* **22**, 153–188.

Ta-Tsien LI [2011]: *Problems and Solutions in Mathematics, Second Edition*, World Scientific, Singapore.

X. LI; R.N. MOHAPATRA [1993]: On the convergence of Lagrange interpolation with equidistant nodes, *Proceedings of the American Mathematical Society* **118**, 1205–1212.

H. LIEBMANN [1899]: Eine neue Eigenschaft der Kugel, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 45–55.

Minghua LIN [2012]: The AM-GM inequality and CBS inequality are equivalent, *The Mathematical Intelligencer* **34**, 6.

J.L. LIONS [1961]: *Equations Différentielles Opérationnelles et Problèmes aux Limites*, Springer, Berlin.

J.L. LIONS [1965]: *Problèmes aux Limites dans les Equations aux Dérivées Partielles*, Presses

de l'Université de Montréal, Montréal, Que.

J.L. LIONS [1969]: *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires*, Dunod, Paris.

J.L. LIONS [1973]: *Perturbations Singulières dans les Problèmes aux Limites et en Contrôle Optimal*, Lecture Notes in Mathematics, Volume 323, Springer, Berlin.

J.L. LIONS; E. MAGENES [1972]: *Non-Homogeneous Boundary Value Problems and Applications, Volume* 1, Springer, Heidelberg (translation of the original French edition, *Problèmes aux Limites non Homogènes et Applications, Volume* 1, Dunod, Paris, 1968).

J.L. LIONS; G. STAMPACCHIA [1967]: Variational inequalities, *Communications on Pure and Applied Mathematics* **20**, 493–519.

P.L. LIONS [1984]: The concentration-compactness principle in the calculus of variations. The locally compact case – Part 1, *Annales de l'Institut Henri Poincaré – Analyse Non Linéaire* **1**, 109–145.

P.L. LIONS [1984]: The concentration-compactness principle in the calculus of variations. The locally compact case – Part 2, *Annales de l'Institut Henri Poincaré – Analyse Non Linéaire* **1**, 223–283.

P.L. LIONS [1985]: The concentration-compactness principle in the calculus of variations. The limit case – Part 1, *Revista Matematica Iberoamericana* **1.1**, 145–201.

P.L. LIONS [1985]: The concentration-compactness principle in the calculus of variations. The limit case – Part 2, *Revista Matematica Iberoamericana* **1.2**, 45–121.

P.L. LIONS [1996]: *Mathematical Topics in Fluid Mechanics, Volume* 1 : *Incompressible Models*, Clarendon Press, Oxford, UK.

J. LIOUVILLE [1850]: Extension au cas des trois dimensions de la question du tracé géographique, Note VI in the Appendix to G. MONGE: *Application de l'Analyse à la Géométrie, Cinquième Edition*, Bachelier, Paris.

G.G. LORENTZ [1986]: *Bernstein Polynomials*, Chelsea, New York.

S. LOZINSKI [1948]: On a class of linear operators, *Doklady Akademii Nauk SSSR* **61**, 193–196 (in Russian).

D.G. LUENBERGER [1969]: *Optimization by Vector Space Methods*, John Wiley, New York.

N. LUSIN [1913]: Sur la convergence des séries trigonométriques de Fourier, *Comptes Rendus de l'Académie des Sciences de Paris* **156**, 1655–1658.

C.R. MACCLUER [2000]: The many proofs and applications of Perron's theorem, *SIAM Review* **42**, 487–498.

E.J. MACSHANE [1934]: Extension of range of functions, *Bulletin of the American Mathematical Society* **40**, 837–842.

E. MAGENES; G. STAMPACCHIA [1958]: I problemi al contorno per le equazioni differenziali di tipo ellitico, *Annali della Scuola Normale Superiore di Pisa* **12**, 247–358.

G. MAINARDI [1856]: Su la teoria generale delle superficie, *Giornale dell' Istituto Lombardo*

9, 385-404.

L. MALIGRANDA [2012]: The AM-GM inequality is equivalent to the Bernoulli inequality, *The Mathematical Intelligencer* **34**, 1–2.

D.H. MALING [1992]: *Coordinate Systems and Map Projections*, Second Edition, Pergamon Press, Oxford.

B. MANIÀ [1934]: Sopra un esempio di Lavrentieff, *Bolletone dell Unione Mathematica Italiana* **13**, 147–153.

C. MARDARE [2003]: On the recovery of a manifold with prescribed metric tensor, *Analysis and Applications* **1**, 433–453.

S. MARDARE [2003]: Inequality of Korn's type on compact surfaces without boundary, *Chinese Annals of Mathematics, Series B*, **24**, 191–204.

S. MARDARE [2005]: On Pfaff systems with L^p coefficients and their applications in differential geometry, *Journal de Mathématiques Pures et Appliquées* **84**, 1659–1692.

S. MARDARE [2007]: On systems of first order linear partial differential equations with L^p coefficients, *Advances in Differential Equations* **12**, 301–360.

S. MARDARE [2008]: On Poincaré and De Rham's theorems, *Revue Roumaine de Mathématiques Pures et Appliquées* **53**, 523–541.

I. MAREK [1970]: Frobenius theory of positive operators: Comparison theorems and applications, *SIAM Journal on Applied Mathematics* **19**, 607–628.

K. MARGUERRE [1939]: Zur Theorie der gekrümmten Platte großer Formänderung, in *Proceedings, Fifth International Congress for Applied Mechanics*, pp. 93–101, John Wiley & Sons, New York, 1939.

A.A. MARKOFF [1889]: Sur une question posée par Mendeleieff, *Izvestia Akademii Nauk SSSR* **62**, 1–24.

J.E. MARSDEN; T.J.R. HUGHES [1999]: *Mathematical Foundations of Elasticity*, Prentice-Hall, Englewood Cliffs, NJ (First Edition: 1983).

R.D. MAUDLIN, editor [1981]: *The Scottish Book—Mathematics from the Scottish Café*, Birkhäuser, Basel.

J. MAWHIN [1979]: *Topological Degree Methods in Nonlinear Boundary Value Problems*, American Mathematical Society, Providence, RI.

J. MAWHIN [1999]: Leray-Schauder degree: A half century of extensions and applications, *Topological Methods in Nonlinear Analysis* **14**, 195–228.

S. MAZUR [1933]: Über konvexe Mengen in linearen normierten Räumen, *Studia Mathematica* **5**, 70–84.

S. MAZUR; S. ULAM [1932]: Sur les transformations isométriques d'espaces vectoriels normés, *Comptes Rendus de l'Académie des Sciences de Paris* **194**, 946–948.

V. MAZ'YA; T. SHAPOSHNIKOVA [1998]: *Jacques Hadamard, a Universal Mathematician*, American Mathematical Society, Providence, RI.

A. McIntosch [1978]: The Toeplitz-Hausdorff theorem and ellipticity conditions, *The American Mathematical Monthly* **85**, 475–477.

W.H. Meeks III; J. Pérez [2011]: The classical theory of minimal surfaces, *Bulletin of the American Mathematical Society* **48**, 325–407.

G.H. Meisters; C. Olech [1963]: Locally one-to-one mappings and a classical theorem on Schlicht functions, *Duke Mathematical Journal* **30**, 63–80.

H.N. Mhaskar; D.V. Pai [2007]: *Fundamentals of Approximation Theory, Revised Edition*, Alpha Science, Oxford, UK (First Edition: 2000).

D.P. Milman [1938]: On some criteria for the regularity of spaces of type (B), *Doklady Akademii Nauk SSSR* **20**, 243–246.

J. Milnor [1965]: *Topology from the Differentiable Viewpoint*, Princeton University Press, Princeton, NJ.

J. Milnor [1978]: Analytic proofs of the "hairy ball theorem" and the Brouwer fixed point theorem, *The American Mathematical Monthly* **85**, 521–524.

H. Minkowski [1896]: *Geometrie der Zahlen*, Leipzig.

G.J. Minty [1962]: Monotone (nonlinear) operators in Hilbert space, *Duke Mathematical Journal* **29**, 341–346.

G.J. Minty [1963]: On a monotonicity method for the solution of nonlinear equations in Banach spaces, *Proceedings of the National Academy of Sciences USA* **50**, 1038–1041.

E.H. Moore [1920]: On the reciprocal of the general algebraic matrix, *Bulletin of the American Mathematical Society* **26**, 394–395.

J.J. Moreau [1970]: Inf-convolution, sous-additivité, convexité des fonctions numériques, *Journal de Mathématiques Pures et Appliquées* **49**, 109–154.

J.J. Moreau [1979]: Duality characterization of strain tensor distributions in an arbitrary open set, *Journal of Mathematical Analysis and Applications* **72**, 760–770.

C.B. Morrey, Jr. [1952]: Quasi-convexity and the lower semicontinuity of multiple integrals, *Pacific Journal of Mathematics* **2**, 25–53.

C.B. Morrey, Jr. [1966]: *Multiple Integrals in the Calculus of Variations*, Springer, Berlin.

P.P. Mosolov; V.P. Mjasnikov [1971]: A proof of Korn's inequality, *Soviet Mathematics Doklady* **12**, 1618–1622.

D. Motreanu; V. Rădulescu [2003]: *Variational and Non-Variational Methods in Nonlinear Analysis and Boundary Value Problems*, Kluwer, Dordrecht.

M.E. Munroe [1953]: *Introduction to Measure and Integration*, Addison-Wesley, Reading, MA.

C. Müntz [1914]: Über den Approximationssatz von Weierstraß, in *H.A. Schwarz Festschrift*, pp. 303–312, Mathematische Abhandlungen, Springer, Berlin.

F. Murat [1978]: Compacité par compensation, *Annali Scuola Normale Superiore de Pisa, Serie* IV, **5**, 489–507.

F. MURAT [1987]: A survey on compensated compactness, in *Contributions to Modern Calculus of Variations* (L. CESARI, editor), pp. 145–183, Longman, Harlow.

L. NACHBIN [1969]: *Topology on Spaces of Holomorphic Mappings*, Springer, Berlin.

M. NAGUMO [1951]: A theory of degree of mapping based on infinitesimal analysis, *American Journal of Mathematics* **73**, 485–496.

J. NASH [1954]: C^1 isometric imbeddings, *Annals of Mathematics* **60**, 383–396.

C.L.M.H. NAVIER [1823]: Mémoire sur les lois du mouvement des fluides, *Mémoires de l'Académie Royale des Sciences de Paris* **6**, 389–416.

J. NEČAS [1962]: Sur une méthode pour résoudre les équations aux dérivées partielles du type elliptique, voisine de la variationnelle, *Annali della Scuola Normale Superiore di Pisa, Classe di Scienze, Serie* III, **16**, 305–326.

J. NEČAS [1965]: *Equations aux Dérivées Partielles*, Presses de l'Université de Montréal, Montréal.

J. NEČAS [1967]: *Les Méthodes Directes en Théorie des Equations Elliptiques*, Masson, Paris and Academia, Praha (English translation: *Direct Methods in the Theory of Elliptic Equations*, Springer, Heidelberg, 2012).

J. NEČAS; I. HLAVÁČEK [1981]: *Mathematical Theory of Elastic and Elasto-Plastic Bodies: An Introduction*, Elsevier, Amsterdam.

P.M. NEUMANN [2011]: *The Mathematical Writings of Évariste Galois*, European Mathematical Society, Zürich.

R.A. NICOLAIDES [1972]: On a class of finite elements generated by Lagrange interpolation, *SIAM Journal on Numerical Analysis* **9**, 435–445.

L. NIRENBERG [1974]: *Topics in Nonlinear Functional Analysis*, Lecture Notes, Courant Institute, New York University, NY (Second Edition: American Mathematical Society, Providence, RI, 1994).

J.A. NITSCHE [1981]: On Korn's second inequality, *RAIRO Analyse Numérique* **15**, 237–248.

J.C.C. NITSCHE [1975]: *Vorlesungen über Minimalflächen*, Springer, Berlin.

B. O'NEILL [2006]: *Elementary Differential Geometry, Revised Second Edition*, Elsevier/ Academic Press, Burlington (First Edition: 1966).

J.T. ODEN; L.F. DEMKOWICZ [2010]: *Applied Functional Analysis, Second Edition*, Chapman & Hall, Boca Raton, FL (First Edition: 1996).

J.M. ORTEGA [1968]: The Newton-Kantorovich theorem, *The American Mathematical Monthly* **75**, 658–660.

J.M. ORTEGA; W.C. RHEINBOLDT [2000]: *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, Philadelphia.

A.M. OSTROWSKI [1954]: On the linear iteration procedures for symmetric matrices, *Rendiconti Lincei – Matematica e Applicazioni* **14**, 140–163.

C. PADOVANI [2000]: On the derivative of some tensor-valued functions, *Journal of Elasticity*

58, 257–268.

R.S. Palais; S. Smale [1964]: A generalized Morse theory, *Bulletin of the American Mathematical Society* **70**, 165–171.

R. Penrose [1955]: A generalized inverse for matrices, *Proceedings of the Cambridge Philosophical Society* **51**, 406–413.

O. Perron [1907]: Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus, *Mathematische Annalen* **64**, 11–76.

O. Perron [1923]: Eine neue Behandlung der Randwertaufgabe für $\Delta u = 0$, *Mathematische Zeitschrift* **18**, 42–54.

B.J. Pettis [1939]: A proof that every uniformly convex space is reflexive, *Duke Mathematical Journal* **5**, 249–253.

R. Phelps [1960]: Uniqueness of Hahn-Banach extensions and unique best approximation, *Transactions of the American Mathematical Society* **95**, 238–255.

E. Picard [1893]: Sur l'application des méthodes d'approximations successives à l'étude de certaines équations différentielles ordinaires, *Journal de Mathématiques Pures et Appliquées* **9**, 217–271.

A. Pietsch [2007]: *History of Banach Spaces and Linear Operators*, Birkhäuser, Boston.

R.B. Platte; L.N. Trefethen; A.B.J. Kuijlaars [2011]: Impossibility of fast stable approximation of analytic functions from equispaced samples, *SIAM Review* **53**, 308–318.

G. Pólya [1933]: Über die Konvergenz von Quadraturverfahren, *Mathematische Zeitschrift* **37**, 264–286.

G. Pólya [1987]: *The Pólya Picture Album: Encounters of a Mathematician*, Birkhäuser, Boston.

A. Pressley [2005]: *Elementary Differential Geometry*, Springer, London.

M. Protter; H. Weinberger [1967]: *Maximum Principles in Differential Equations*, Prentice-Hall, Englewood Cliffs, NJ.

P. Pucci; J. Serrin [2007]: *The Maximum Principle*, Birkhäuser, Basel.

P. Rabier [1979]: Résultats d'existence dans des modèles non linéaires de plaques, *Comptes Rendus de l'Académie des Sciences de Paris, Série A*, **289**, 515–518.

R. Rado [1956]: Note on generalized inverses of matrices, *Proceedings of the Cambridge Philosophical Society* **52**, 600–601.

T. Rado [1930]: The problem of the least area and the problem of Plateau, *Mathematische Zeitschrift* **32**, 763–796.

T. Rado; P.V. Reichelderfer [1955]: *Continuous Transformations in Analysis*, Springer, Berlin.

I.K. Rana [2002]: *An Introduction to Measure and Integration, Second Edition*, Graduate Studies in Mathematics, Volume 45, American Mathematical Society, Providence, RI.

P.A. Raviart; J.M. Thomas [1983]: *Introduction à l'Analyse Numérique des Equations aux*

Dérivées Partielles, Masson, Paris.

E. Reich [1949]: On the convergence of the classical iterative method of solving linear simultaneous equations, *Annals of Mathematical Statistics* **20**, 448–451.

C. Reid [1970]: *Hilbert—With an Appreciation of Hilbert's Mathematical Work by Hermann Weyl*, Springer, New York.

C. Reid [1976]: *Courant in Göttingen and New York—The Story of an Improbable Mathematician*, Springer, New York.

F. Rellich [1930]: Ein Satz über mittlere Konvergenz, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, 30–35.

Y.G. Reshetnyak [1967]: Liouville's theory on conformal mappings under minimal regularity assumptions, *Siberian Mathematical Journal* **8**, 69–85.

G. de Rham [1955]: *Variétés Différentiables*, Hermann, Paris.

W.C. Rheinboldt [1968]: A unified convergence theory for a class of iterative processes, *SIAM Journal on Numerical Analysis* **5**, 42–63.

F. Riesz [1907]: Sur les systèmes orthogonaux de fonctions, *Comptes Rendus de l'Académie des Sciences* **144**, 615–619.

F. Riesz [1907]: Sur une espèce de géométrie analytique des systèmes de fonctions sommables, *Comptes Rendus de l'Académie des Sciences de Paris* **144**, 1409–1411.

F. Riesz; B. Sz.-Nagy [1955]: *Leçons d'Analyse Fonctionnelle, Troisième Edition*, Gauthier-Villars, Paris, and Akadémiai Kiadó, Budapest (English translation: *Functional Analysis*, Dover, New York, 1990).

J.E. Roberts; J.M. Thomas [1991]: Mixed and hybrid methods, in *Handbook of Numerical Analysis, Volume* II (P.G. Ciarlet & J.L. Lions, editors), pp. 523–639, North-Holland, Amsterdam.

H.L. Royden [1963]: *Real Analysis*, MacMillan, New York (Third Edition: 1988).

W. Rudin [1966]: *Real and Complex Analysis*, McGraw-Hill, New York (Third Edition: 1987).

W. Rudin [1973]: *Functional Analysis*, McGraw-Hill, New York (Second Edition: 1991).

W. Rudin [1997]: *The Way I Remember It*, American Mathematical Society, Providence, RI.

H. Samelson [2001]: Differential forms, the early days; or the stories of Deahna's theorem and of Volterra's theorem, *American Mathematical Monthly* **108**, 552–530.

A. Sard [1942]: The measure of the critical values of differential maps, *Bulletin of the American Mathematical Society* **48**, 883–890.

F. Sauvigny [2006a]: *Partial Differential Equations* 1: *Foundations and Integral Representations*, Springer, Berlin.

F. Sauvigny [2006b]: *Partial Differential Equations* 2: *Functional Analytic Methods*, Springer, Berlin.

G.M. Scarpello; D. Ritelli [2002]: A historical outline of the theorem of implicit functions, *Divulgaciones Matemáticas* **10**, 171–180.

H. Schäfer [1955]: Über die Methode der a priori Schranken, *Mathematische Annalen* **129**, 415–416.

J. Schauder [1930]: Der Fixpunktsatz in Funktionalräumen, *Studia Mathematica* **2**, 171–180.

J. Schauder [1934]: Über lineare elliptische Differentialgleichungen zweiter Ordnung, *Mathematische Zeitschrift* **38**, 257–282.

M. Schechter [1971]: *Principles of Functional Analysis, First Edition*, Graduate Studies in Mathematics, Volume 36, American Mathematical Society, Providence, RI (Second Edition: 2002).

H. Schlichtkrull [2012]: *Differential Manifolds*, Lecture Notes for Geometry 2, available online at www.math.ku.dk/~jakobsen/geom2/manusgeom2.pdf.

E. Schmidt [1907]: Zur Theorie der linearen und nichtlinearen Integralgleichungen. 1. Teil: Entwicklung willkürlicher Funktionen nach Systemen vorgeschriebener, *Mathematische Annalen* **63**, 433–476.

J. Schwartz [1969]: *Nonlinear Functional Analysis*, Gordon and Breach, New York.

L. Schwartz [1965]: *Méthodes Mathématiques pour les Sciences Physiques*, Hermann, Paris (English translation: *Mathematics for the Physical Sciences*, Dover, New York, 2008).

L. Schwartz [1966]: *Théorie des Distributions*, Hermann, Paris.

L. Schwartz [1970]: *Analyse: Deuxième Partie: Topologie Générale et Analyse Fonctionnelle*, Hermann, Paris.

L. Schwartz [1991]: *Analyse* I: *Théorie des Ensembles et Topologie*, Hermann, Paris.

L. Schwartz [1992]: *Analyse* II: *Calcul Différentiel et Equations Différentielles*, Hermann, Paris.

L. Schwartz [1993a]: *Analyse* III: *Calcul Intégral*, Hermann, Paris.

L. Schwartz [1993b]: *Analyse* IV: *Applications de la Théorie de la Mesure*, Hermann, Paris.

L. Schwartz [2001]: *A Mathematician Grappling with His Century*, Birkhäuser, Basel (translation of the original French edition, *Un Mathématicien aux Prises avec le Siècle*, Odile Jacob, Paris, 1997).

H.A. Schwarz [1885]: Über ein Flächen kleinsten Flächeninhalts betreffendes Problem der Variationsrechnung, *Acta Societatis Scientiarum Fennicae* **15**, 315–362.

D. Serre [2010]: *Matrices, Second Edition*, Springer, Heidelberg (translated from the original French edition, *Matrices*, Springer, New York, 2002).

R.T. Shield [1973]: The rotation associated with large strains, *SIAM Journal on Applied Mathematics* **25**, 483–491.

A. Signorini: Sopra alcune questioni di elastostatica, *Atti della Società Italiana per il Progresso della Scienza* (1933).

J.G. SIMMONDS [1994]: *A Brief on Tensor Analysis, Second Edition*, Springer, Berlin (First Edition: 1982).

M. SION [1958]: On general mini-max theorems, *Pacific Journal of Mathematics* **8**, 171–176.

S. SLICARU [1998]: On the ellipticity of the middle surface of a shell and its application to the asymptotic analysis of "membrane shells," *Journal of Elasticity* **46**, 33–42.

K.T. SMITH [1983]: *Primer of Modern Analysis*, *Second Edition*, Springer, New York (First Edition: 1971, Bogden & Quigley, Tarrytown-on-Hudson, NY).

S.J. SMITH [2006]: Lebesgue constants in polynomial interpolation, *Annales Mathematicae et Informaticae* **33**, 109–123.

V.L. ŠMULIAN [1940]: Über lineare topologische Räume, *Mathematiceskii Sbornik, N.S.* **49**, 425–448.

J.P. SNYDER [1993]: *Flattening the Earth: Two Thousand Years of Map Projection*, University of Chicago Press, Chicago.

S.L. SOBOLEV [1938]: On a theorem of functional analysis, *Matematicheskii Sbornik* **46**, 471–496.

S.L. SOBOLEV [1950]: *Applications of Functional Analysis in Mathematical Physics*, Leningrad (in Russian; English translation: American Mathematical Society, Providence, RI, 1963).

V.A. SOLONNIKOV [1982]: On the Stokes equations in domains with non-smooth boundaries and on viscous incompressible flow with a free surface, in *Nonlinear Partial Differential Equations and Their Applications* (H. BREZIS & J.L. LIONS, editors), pp. 340–423, Pitman, Boston.

G.A. SOUKHOMLINOFF [1938]: Über Fortsetzung von linearen Funktionalen in linearen komplexen Räumen und linearen Quaternionräumen, *Mathematiceskii Sbornik* **3**, 353–358.

M. SPIVAK [1999]: *A Comprehensive Introduction to Differential Geometry, Volumes* I *to* V, *Third Edition*, Publish or Perish, Berkeley, CA.

I. STAKGOLD [1998]: *Green's Functions and Boundary Value Problems, Second Edition*, John Wiley, New York (First Edition: 1979).

G. STAMPACCHIA [1964]: Formes bilinéaires coercitives sur les ensembles convexes, *Comptes Rendus de l'Académie des Sciences de Paris Série A*, **258**, 4413–4416.

G. STAMPACCHIA [1965]: *Equations Elliptiques du Second Ordre à Coefficients Discontinus*, Presses de l'Université de Montréal, Montréal, Que.

G. STAMPACCHIA [1965]: Le problème de Dirichlet pour les équations elliptiques du second ordre à coefficients discontinus, *Annales de l'Institut Fourier* (*Grenoble*) **15**, 189–258.

E.M. STEIN [1970]: *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, NJ.

E.M. STEIN; R. SHAKARCHI [2005]: *Real Analysis: Measure Theory, Integration and Hilbert Spaces*, Princeton Lectures on Analysis, Volume III, Princeton University Press, Princeton, NJ.

E.M. STEIN; R. SHAKARCHI [2011]: *Functional Analysis: Introduction to Further Topics in Analysis*, Princeton University Press, Princeton, NJ.

H. STEINLEIN [1979]: Two results of J. Dugundji about extensions of maps and retractions, *Proceedings of the American Mathematical Society* **77**, 290–298.

R.A. STEPHENSON [1980]: On the uniqueness of the square-root of a symmetric, positive-definite tensor, *Journal of Elasticity* **10**, 213–214.

G.W. STEWART [1969]: On the continuity of the generalized inverse, *SIAM Journal on Applied Mathematics* **17**, 33–45.

J.J. STOKER [1969]: *Differential Geometry*, John Wiley, New York.

G.G. STOKES [1845]: On the theories of the internal friction of fluids in motion, *Transactions of the Cambridge Philosophical Society* **8**, 287–305.

M.H. STONE [1948]: The generalized Weierstrass approximation theorem, *Mathematics Magazine* **21**, 167–183 and 237–254.

G. STRANG [1976]: *Linear Algebra and Its Applications*, Academic Press, New York.

G. STRANG [2009]: *Introduction to Linear Algebra, Fourth Edition*, Wellesley Cambridge Press, UK.

M. STRUWE [1990]: *Variational Methods—Applications to Nonlinear Partial Differential Equations and Hamiltonian Systems*, Springer, Berlin.

A. STUBHAUG [2000]: *Niels Henrik Abel and his Times—Called Too Soon by Flames Afar*, Springer, Heidelberg (translated from the Norwegian).

R.H. SZCZARBA [1970]: On isometric immersions of Riemannian manifolds in Euclidean space, *Boletim da Sociedade Brasileira de Matemática* **1**, 31–45.

G. SZEGŐ [1975]: *Orthogonal Polynomials, Fourth Edition*, American Mathematical Society, Providence, RI (First Edition: 1939).

M. SZOPOS [2005]: On the recovery and continuity of a submanifold with boundary, *Analysis and Applications* **3**, 119–143.

L. TARTAR [1978]: *Topics in Nonlinear Analysis*, Publications Mathématiques d'Orsay No. 78.13, Université de Paris-Sud, Orsay.

L. TARTAR [1979]: Compensated compactness and partial differential equations, in *Nonlinear Analysis and Mechanics, Heriot-Watt Symposium, Volume* IV (R. J. KNOPS, editor), pp. 136–212, Pitman, Boston.

L. TARTAR [1983]: The compensated compactness method applied to systems of conservation laws, in *Systems of Nonlinear Partial Differential Equations* (J.M. BALL, editor), pp. 263–285, Reidel, Dordrecht.

L. TARTAR [2006]: *An Introduction to Navier-Stokes Equation and Oceanography*, Springer, Berlin.

L. TARTAR [2007]: *An Introduction to Sobolev Spaces and Interpolation Spaces*, Springer, Berlin.

L. TARTAR [2009]: *The General Theory of Homogenization: A Personalized Introduction*, Springer, Berlin.

A.E. TAYLOR [1939]: The extension of linear functionals, *Duke Mathematical Journal* **5**, 538–547.

A.E. TAYLOR [1958]: *Introduction to Functional Analysis*, John Wiley, New York.

A.E. TAYLOR [1965]: *General Theory of Functions and Integration*, Blaisdell, Waltham.

A.E. TAYLOR; D.C. LAY [1980]: *Introduction to Functional Analysis, Second Edition*, John Wiley, New York.

M.E. TAYLOR [1996a]: *Partial Differential Equations* I: *Basic Theory*, Springer, New York.

M.E. TAYLOR [1996b]: *Partial Differential Equations* II: *Qualitative Studies of Linear Equations*, Springer, New York.

M.E. TAYLOR [1996c]: *Partial Differential Equations* III: *Nonlinear Equations*, Springer, New York.

R. TEMAM [1971]: Solutions généralisées de certaines équations du type hypersurfaces minima, *Archive for Rational Mechanics and Analysis* **44**, 121–156.

R. TEMAM [1977]: *Navier–Stokes Equations*, North-Holland, Amsterdam.

R. TEMAM [1995]: *Navier–Stokes Equations and Nonlinear Functional Analysis, Second Edition*, SIAM, Philadelphia.

K. TENENBLAT [1971]: On isometric immersions of Riemannian manifolds, *Boletim da Sociedade Brasileira de Matemática* **2**, 23–36.

T.Y. THOMAS [1934]: Systems of total differential equations defined over simply connected domains, *Annals of Mathematics* **35**, 730–734.

T.W. TING [1974]: St. Venant's compatibility conditions, *Tensors, N.S.* **28**, 5–12.

K. TINTAREV; K.-H. FIESELER [2007]: *Concentration Compactness. Functional-Analytic Grounds and Applications*, Imperial College Press, London.

O. TOEPLITZ [1918]: Das algebraische Analogon zu einem Satze von Fejér, *Mathematische Zeitschrift* **2**, 187–197.

L. TONELLI [1920]: La semicontinuità nel calcolo delle variazioni, *Rendiconti del Circolo Matematico di Palermo* **44**, 167–249.

A. TYCHONOFF [1930]: Über die topologische Erweiterung von Räumen, *Mathematische Annalen* **102**, 544–561.

S.M. ULAM [1976]: *Adventures of a Mathematician*, reprinted and expanded by University of California Press, Berkeley, 1991.

H. UZAWA [1958]: Iterative methods for concave programming, in *Studies in Linear and Nonlinear Programming* (K.J. ARROW, L. HURWICZ, & H. UZAWA, editors), pp. 154–165, Stanford University Press, Stanford, CA.

M.M. VAINBERG [1952]: Some questions of differential calculus in linear spaces, *Uspehi Matematicheskii Nauk* (*New Series*) **7**, 55–102 (in Russian).

T. VALENT [1988]: *Boundary Value Problems of Finite Elasticity—Local Theorems on Existence, Uniqueness, and Analytic Dependence on Data*, Springer, New York.

C.J. DE LA VALLÉE POUSSIN [1910]: Sur les polynômes d'approximation et la représentation approchée d'un angle, *Académie Royale de Belgique, Bulletins de la Classe des Sciences* **12**.

C. VALLÉE; D. FORTUNÉ [1976]: Compatibility equations in shell theory, *International Journal of Engineering Science* **34**, 495–499.

R.S. VARGA [1962]: *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ.

K. VO-KHAC [1972a]: *Distributions—Analyse de Fourier—Opérateurs aux Dérivées Partielles, Volume* 1, Vuibert, Paris.

K. VO-KHAC [1972b]: *Distributions—Analyse de Fourier—Opérateurs aux Dérivées Partielles, Volume* 2, Vuibert, Paris.

V. VOLTERRA [1907]: Sur l'équilibre des corps élastiques multiplement connexes, *Annales de l'Ecole Normale* **24**, 401–517.

E.V. VORONOVSKAJA [1932]: Détermination de la forme asymptotique de l'approximation des fonctions par les polynômes de M. Bernstein, *Doklady Akademii Nauk SSSR* **4**, 79–85.

K. WEIERSTRASS [1872]: Über continuirliche Functionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten besitzen, *Königliche Akademie der Wissenschaften*.

K. WEIERSTRASS [1885]: Über die analytische Darstellbarkeit sogenannter willkürlicher Funktionen einer reellen Veränderlichen, *Sitzungsberichte der Akademie zu Berlin*, 633–639 and 789–805.

J. WEINGARTEN [1861]: Über eine Klasse auf einander abwickelbarer Flächen, *Journal für Reine und Angewandte Mathematik* **59**, 382–393.

R.S. WESTFALL [1980]: *Never at Rest: A Biography of Isaac Newton*, Cambridge University Press, Cambridge, UK.

H. WEYL [1940]: The method of orthogonal projection in potential theory, *Duke Mathematical Journal* **7**, 414–444.

R. WHITLEY [1967]: An elementary proof of the Eberlein-Šmulian theorem, *Mathematische Annalen* **172**, 116–118.

H. WHITNEY [1934]: Analytic extensions of differentiable functions defined in closed sets, *Transactions of the American Mathematical Society* **36**, 63–89.

R. WONG [2010]: *Lecture Notes on Applied Analysis*, World Scientific, Singapore.

Q. YANG; J.P. SNYDER; W.R. TOBLER [2000]: *Map Projection Transformation—Principle and Applications*, Taylor and Francis, London.

W.H. YOUNG [1910]: *The Fundamental Theorems of the Differentiable Calculus*, Cambridge University Press, Cambridge, UK.

K. YOSIDA [1966]: *Functional Analysis, First Edition*, Springer, Berlin (Reprint of the Sixth

Edition: 1980).

G. ZAMPIERI [1992]: Diffeomorphisms with Banach space domains, *Nonlinear Analysis, Theory, Methods & Applications* **19**, 923–932.

F. ZARANTONELLO [1960]: Solving functional equations by contractive averaging, *Mathematics Research Center Report No.* 160, University of wisconsin Madison, Madison, WI.

E. ZEIDLER [1985]: *Nonlinear Functional Analysis and Its Applications, Volume* III: *Variational Methods and Optimization*, Springer, New York.

E. ZEIDLER [1986]: *Nonlinear Functional Analysis and Its Applications, Volume* I: *Fixed-Point Theorems*, Springer, Berlin.

E. ZEIDLER [1990a]: *Nonlinear Functional Analysis and Its Applications, Volume* IIa: *Linear Monotone Operators*, Springer, New York.

E. ZEIDLER [1990b]: *Nonlinear Functional Analysis and Its Applications, Volume* IIb: *Fixed-Point Theorems*, Springer, New York.

E. ZEIDLER [1995a]: *Applied Functional Analysis: Main Principles and Their Applications*, Springer, New York.

E. ZEIDLER [1995b]: *Applied Functional Analysis: Applications of Mathematical Physics*, Springer, New York.

E. ZERMELO [1904]: Beweis dass jede Menge wohlgeordnet werden kann, *Mathematische Annalen* **LIX**, 514–516.

M. ZLÁMAL [1968]: On the finite element method, *Numerische Mathematik* **12**, 394–409.

主 要 符 号

集合、映射、序列

$\varnothing$: 空集.

$A \subset B$: A 包含在 B 中.

$A \subsetneqq B$: A 严格包含在 B 中.

$A \bigcup B$: A 和 B 的并集.

$A \bigcap B$: A 和 B 的交集.

$A \times B$: A 和 B 的积.

$\bigcup_{i \in I} A_i$: 一族集合 $(A_i)_{i \in I}$ 的并集.

$\bigsqcup_{i \in I} A_i$: 一族集合 $(A_i)_{i \in I}$ 的不相交的并集.

$\bigcap_{i \in I} A_i$: 一族集合 $(A_i)_{i \in I}$ 的交集.

$\Pi_{i \in I} A_i$: 一族集合 $(A_i)_{i \in I}$ 的积.

$X - A = \{y \in X; y \notin A\}$: 子集 $A \in X$ 的余集.

$\mathbb{N} = \{0, 1, 2, \cdots\}$: 自然数集.

$\mathbb{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$: 整数集.

$\mathbb{Q}$: 有理数集.

$\mathbb{R}$: 实数集.

$\{-\infty\} \bigcup \mathbb{R} \bigcup \{\infty\}$: 扩充的实数集.

$\mathbb{C}$: 复数集.

$\mathbb{K} = \mathbb{R}$ 或 $\mathbb{C}$: 数集.

$\bar{z} : z \in \mathbb{C}$ 的复共轭.

$\mathrm{Re}\, z$ 及 $\mathrm{Im}\, z : z \in \mathbb{C}$ 的实部及虚部.

δ_{ij}, δ_i^j 或 δ^{ij}: Kronecker 符号 ($\delta_{ij} = 1$ 若 $i = j, \delta_{ij} = 0$ 若 $i \neq j$).

$\mathfrak{S}_n : \{1, 2, \cdots, n\}$ 的所有置换的集合.

$\bar{A}$: 集合 A 的闭包.

$\overset{\circ}{A}$ 或 $\mathrm{int}\, A$: 集合 A 的内部.

∂A: 集合 A 的边界.

$\mathrm{card}\, A$: 集合 A 的基数.

$f : X \to Y$ 或 $f : x \in X \to f(x) \in Y$: X 到 Y 的映射或函数.

$g \circ f$: f 和 g 的复合.

$f|_A$: 映射 f 在集合 A 上的限制.

$f(\cdot, b)$: 部分映射 $x \to f(x, b)$.

$f(A) = \{y \in Y; y = f(x)$ 对某个 $x \in X\}$: 子集 $A \subset X$ 在映射 $f : X \to Y$ 下的像 (也表示为 $\mathrm{Im}(A)$ 若 A 是线性的).

$f^{-1}(B) = \{x \in X; f(x) \in B\}$: 子集 $B \subset Y$ 在映射 $f : X \to Y$ 下的逆像.

f^{-1}: 双射的逆映射.

$\mathrm{supp}\, f = \overline{\{x \in X; f(x) \neq 0\}}$: 函数 $f : X \to \mathbb{R}$ 的支集.

id 或 id_X: 集合 X 的恒等映射.

$\mathrm{sgn}\, \alpha = 1$ 若 $\alpha > 0, \mathrm{sgn}\, \alpha = 0$ 若 $\alpha = 0, \mathrm{sgn}\, \alpha = -1$ 若 $\alpha < 0$.

$\deg(f, \Omega, b)$: 映射 $f \in \mathcal{C}(\overline{\Omega}; \mathbb{R}^n)$ 在点 $b \notin f(\partial\Omega)$ 的 Brouwer 拓扑度 (在此, Ω 是 $\mathbb{R}^n$ 的有界开子集).

$(x_k)_{k=\ell}^{\infty}$ 或 (x_k) 若 $\ell = 0$ 或者 $\ell = 1$: 元素 $x_\ell, x_{\ell+1}, \cdots, x_k, \cdots$ 的序列.

$(x_{\sigma(k)})_{k=1}^{\infty} : (x_k)_{k=1}^{\infty}$ 的子序列 (其中 σ 表示集合 $\{1, 2, \cdots\}$ 到其自身的任一严格递增映射).

$x = \lim_{k\to\infty} x_k$, $x_k \underset{k\to\infty}{\longrightarrow} x$ 或 $x_k \to x$ 当 $k \to \infty$ 时: 序列 (x_k) 收敛, 且其极限是 x.

$\liminf_{k\to\infty} x_k, \limsup_{k\to\infty} x_k$: 在集合 $\{-\infty\} \bigcup \mathbb{R} \bigcup \{\infty\}$ 中数的序列 (x_k) 的下极限, 上极限.

在不至于引起混淆的情况, 为简洁起见, 有时省去符号 “$k \to \infty$” (即 $x = \lim x_k, x = \limsup x_k, x_k \to x$ 等).

$x \to a^+$: 实数 $x > a$ 收敛于 $a \in \mathbb{R}$.

$x \to a^-$: 实数 $x < a$ 收敛于 $a \in \mathbb{R}$.

dx, meas 或 dx-meas: n 维 Lebesgue 测度.

向量空间

$X = Y \oplus Z$: X 是子空间 Y 和 Z 的直和.

$[a, b] = \{ta + (1-t)b; 0 \leqslant t \leqslant 1\}$: 端点为 a 和 b 的闭区间.

$]a, b[= \{ta + (1-t)b; 0 < t < 1\}$: 端点为 a 和 b 的开区间.

I: 向量空间的恒等映射.

$\text{Ker } A = \{x \in X; Ax = 0\}$: 线性算子 $A : X \to Y$ 的核.

$\text{Im } A = \{y \in Y; y = Ax$ 对某个 $x \in X\}$: 空间 X 在线性算子 $A : X \to Y$ 映射下的像 (也表为 $A(X)$).

$(X, \|\cdot\|)$: 装备范数 $\|\cdot\|$ 的向量空间 X.

$\|\cdot\|_X$: 向量空间 X 中的范数.

$\|\cdot\|_p$: 空间 ℓ^p 的范数, $1 \leqslant p \leqslant \infty$.

$(X, (\cdot, \cdot))$: 装备内积 $(\cdot, \cdot)$ 的 Hilbert 空间 X.

$|\cdot|$: $\mathbb{R}^n$ 中的欧氏范数.

$|\cdot|$: 矩阵的从属于欧氏范数的算子范数.

$B(a; x) = \{x \in X; \|x - a\| < r\}$: 球心在 a, 半径为 r 的开球.

$\text{co } A$: 集合 A 的凸包.

$\overline{\text{co}}\, A$: 集合 A 的闭凸包.

$\mathcal{L}(X; Y)$: 所有从赋范向量空间 X 到赋范向量空间 Y 的连续线性映射的空间.

$\mathcal{L}(X) = \mathcal{L}(X; X)$.

$X' = \mathcal{L}(X; \mathbb{K})$: $\mathbb{K}$ 上的赋范向量空间 X 的对偶 (空间).

${}_{X'}\langle x', x\rangle_X = x'(X)$ 对任意 $x' \in X'$ 和 $x \in X$.

$X'' = \mathcal{L}(X'; \mathbb{K})$: $\mathbb{K}$ 上的赋范向量空间 X 的双对偶 (空间).

$A' \in \mathcal{L}(Y'; X)$: 线性算子 $A \in \mathcal{L}(X; Y)$ 的对偶 (算子).

$A^* \in \mathcal{L}(Y; X)$: 当 X 和 Y 是 Hilbert 空间时, $A \in \mathcal{L}(X; Y)$ 的伴随 (算子).

$\mathcal{L}_k(X_1, X_2, \cdots, X_k; Y)$ 或 $\mathcal{L}_k(X; Y)$ 若 $X := X_1 = X_2 = \cdots = X_k$: 所有从赋范向量空间的积 $X_1 \times X_2 \times \cdots \times X_k$ 到赋范向量空间 Y 的连续 k 线性映射的空间, $k \geqslant 2$.

X/Y: 向量空间 X 关于 X 的向量子空间 Y 的商空间.

$X \hookrightarrow Y$: X 包含在 Y 内并伴有一个连续内射.

$X \Subset Y$: X 包含在 Y 内并伴有一个紧内射.

$A^\perp = \{y \in X; (y, x) = 0$ 对所有 $x \in A\}$: Hilbert 空间 $(X, (\cdot, \cdot))$ 的子集 A 的正交补.

$x_k \to x$ 或 $x = \lim x_k$: 在 $(X; \|\cdot\|)$ 中强收敛, 即 $\lim \|x_k - x\|_X = 0$.

$x_k \rightharpoonup x$: 在 X 中弱收敛, 即 $\lim x'(x_k) = x'(x)$ 对所有 $x' \in X'$.

$x_k \overset{*}{\rightharpoonup} x$: 在 X' 中弱 * 收敛, 即 $\lim x'_k(x) = x'(x)$ 对所有 $x \in X$.

函数空间

$\mathcal{P}_n$: 所有次数 $\leqslant n$ 的实多项式的空间.

$\mathcal{P} := \bigcup_{n=1}^{\infty} P_n$: 所有实多项式空间.

$\mathcal{P}_n[a,b] = \{p|_{[a,b]}; p \in \mathcal{P}_n\}$.

$\mathcal{P}[a,b] = \{p|_{[a,b]}; p \in \mathcal{P}\}$.

$\mathcal{P}([a,b];\mathbb{C}) := \{p|_{[a,b]};\ p$ 是复系数多项式$\}$.

$\mathcal{Q}_n[0,2\pi]$: 所有阶数 $\leqslant n$ 的实 2π 周期三角多项式的空间.

$\mathcal{Q}[0,2\pi] = \bigcup_{n=0}^{\infty} \mathcal{Q}_n[0,\pi]$: 所有实 2π 周期三角多项式的空间.

$\mathcal{Q}_n([0,2\pi];\mathbb{C})$: 所有阶数 $\leqslant n$ 的复 2π 周期三角多项式的空间.

$\mathcal{Q}([0,2\pi];\mathbb{C}) = \bigcup_{n=0}^{\infty} \mathcal{Q}_n([0,2\pi];\mathbb{C})$: 所有复 2π 周期三角多项式的空间.

$\mathcal{C}(X;Y)$: 所有从拓扑空间 X 到拓扑空间 Y 的连续映射的集合.

$\mathcal{C}(X) = \mathcal{C}(X;\mathbb{R})$.

$\mathcal{C}[a,b] = \mathcal{C}([a,b];\mathbb{R})$.

$\mathcal{C}_{\text{per}}[0,2\pi] = \{g \in \mathcal{C}[0,2\pi]; g(0) = g(2\pi)\}$.

$\mathcal{C}^m(\Omega;Y)$: 所有从一个赋范向量空间的开子集 Ω 到赋范向量空间 Y 的 m 次连续可微映射的空间, $1 \leqslant m \leqslant \infty$.

$\mathcal{C}^m(\Omega) = \mathcal{C}^m(\Omega;\mathbb{R})$.

$\mathcal{C}^m(\overline{\Omega})$, 其中 Ω 是 $\mathbb{R}^n$ 的有界开子集, $1 \leqslant m \leqslant \infty$: 所有满足如下条件的函数 $v \in \mathcal{C}^m(\Omega)$ 的空间, 即对每个重指标 $\boldsymbol{\alpha}, |\boldsymbol{\alpha}| \leqslant m$, 存在函数 $v^{\boldsymbol{\alpha}} \in \mathcal{C}^0(\overline{\Omega})$ 使得 $v^{\boldsymbol{\alpha}}|_{\Omega} = \partial^{\boldsymbol{\alpha}} v$.

$\|v\|_{\mathcal{C}^m(\overline{\Omega})} = \max_{|\alpha| \leqslant m} \sup_{x \in \overline{\Omega}} |v^{\alpha}(x)|$.

$\mathcal{C}^m[a,b] = \{f|_{[a,b]}; f \in \mathcal{C}^m(\mathbb{R})\}$.

在下面, Ω 是 $\mathbb{R}^n$ 的开子集, 或是 $\mathbb{R}^n$ 中的区域.

$\mathcal{D}(\Omega) = \{v \in \mathcal{C}^{\infty}(\Omega); \operatorname{supp} v$ 是 Ω 的紧子集$\}$.

$\mathcal{D}'(\Omega) : \Omega$ 上分布的空间.

$L^p(\Omega), L^p(\Omega;\mathbb{C}), 1 \leqslant p \leqslant \infty : dx$ 几乎处处等于一个满足 $\|v\|_{0,p,\Omega} < \infty$ 的函数 v 的等价类构成的空间; 后者表示满足同样条件的复值函数的等价类空间.

$\|v\|_{0,\infty,\Omega} = \inf\{\alpha \geqslant 0; dx\text{-meas}\{x \in \Omega; |v(x)| \geqslant \alpha\} = 0\}$ 若 $p = \infty$.

$\|v\|_{0,p,\Omega} = \left\{ \int_{\Omega} |v(x)|^p dx \right\}^{1/p} < \infty$ 若 $1 \leqslant p < \infty$.

$\|v\|_{0,\Omega} = \|v\|_{0,2,\Omega}$.

$L^p(a,b) = L^p(\Omega),\ \Omega =]a,b[\subset \mathbb{R}$.

$L^p(\Gamma), 1 \leqslant p < \infty$, 其中 $\Gamma = \partial\Omega : d\Gamma$ 几乎处处等于满足 $\int_\Gamma |v|^p d\Gamma < \infty$ 的函数的等价类空间.

$\|v\|_{L^p(\Gamma)} = \left\{\int_\Gamma |v|^p d\Gamma\right\}^{1/p} < \infty, 1 \leqslant p < \infty.$

$W^{m,p}(\Omega) = \{v \in L^p(\Omega); \partial^\alpha v \in L^p(\Omega)$ 对所有 $|\alpha| \leqslant m\}, 1 \leqslant m, 1 \leqslant p \leqslant \infty.$

$W_0^{m,p}(\Omega) : \mathcal{D}(\Omega)$ 在 $W^{m,p}(\Omega)$ 中的闭包, $1 \leqslant m, 1 \leqslant p < \infty$.

$\|v\|_{m,p,\Omega} = \left\{\int_\Omega \sum_{|\boldsymbol{\alpha} \leqslant m|} |\partial^{\boldsymbol{\alpha}} v|^p dx\right\}^{1/p}, \quad 1 \leqslant m, 1 \leqslant p < \infty.$

$\|v\|_{m,\infty,\Omega} = \max_{|\boldsymbol{\alpha}| \leqslant m} |\partial^{\boldsymbol{\alpha}} v|_{0,\infty,\Omega}.$

$|v|_{m,p,\Omega} = \left\{\int_\Omega \sum_{|\boldsymbol{\alpha}| = m} |\partial^{\boldsymbol{\alpha}} v|^p dx\right\}^{1/p}, \quad 1 \leqslant m, 1 \leqslant p < \infty.$

$|v|_{m,\infty,\Omega} = \max_{|\boldsymbol{\alpha}| = m} |\partial^{\boldsymbol{\alpha}} v|_{0,\infty,\Omega}, \quad 1 \leqslant m.$

$H^m(\Omega) = W^{m,2}(\Omega), \quad 1 \leqslant m.$

$H_0^m(\Omega) = W_0^{m,2}(\Omega), \quad 1 \leqslant m.$

$\|v\|_{m,\Omega} = \|v\|_{m,2,\Omega}, \quad 1 \leqslant m.$

$|v|_{m,\Omega} = |v|_{m,2,\Omega}, \quad 1 \leqslant m.$

$\mathrm{tr} \in \mathcal{L}(W^{1,p}(\Omega); L^q(\Gamma))$: 从 Sobolev 空间 $W^{1,p}(\Omega), 1 \leqslant p < \infty$, 到空间 $L^q(\Gamma)$ 的迹算子 ($\mathrm{tr}\,\boldsymbol{A}$ 也表示矩阵 $\boldsymbol{A}$ 的迹).

如果以 $V(\Omega)$ 表示定义在 Ω 上的实值函数的空间, 那么 $\boldsymbol{V}(\Omega)$ 和 $\mathbb{V}(\Omega)$ 则分别表示相应的向量值空间和对称张量值空间, 其分量或元素属于 $V(\Omega)$; 例如

$\boldsymbol{W}^{1,p}(\Omega) = \{\boldsymbol{v} = (v_i); v_i \in W^{1,p}(\Omega)\},$

$\mathbb{L}^2(\Omega) = \{\boldsymbol{\sigma} = (\sigma_{ij}); \sigma_{ij} = \sigma_{ji} \in L^2(\Omega)\},$

等等, 其相应的范数或半范数均用同样的符号表示; 例如

$\|\boldsymbol{v}\|_{1,p,\Omega} = \left(\sum_i \|v_i\|_{1,p,\Omega}^p\right)^{1/p}$ 对每个 $\boldsymbol{v} = (v_i) \in \boldsymbol{W}^{1,p}(\Omega)$,

$\|\boldsymbol{\sigma}\|_{0,\Omega} = \left(\sum_{i,j} \|\sigma_{ij}\|_{0,\Omega}^2\right)^{1/2}$ 对每个 $\boldsymbol{\sigma} = (\sigma_{ij}) \in \mathbb{L}^2(\Omega)$, 等等.

微分学

下面, X 和 Y 表示赋向量空间, Ω 是 X 的开子集, 而 f 是从 Ω 到 Y 的映射.

$f'(a) \in \mathcal{L}(X;Y)$: f 在 $a \in \Omega$ 处的 Fréchet 导数.

$\frac{df}{dx}(a) = f'(a)$ 若 $X = \mathbb{R}$.

$\partial_j f(a) \in \mathcal{L}(X_j;Y)$ 或 $\dfrac{\partial f}{\partial x_j}(a)$: f 在 a 处的第 j 个偏导数, 当 $X = \prod_{j=1}^n X_j$ 时.

$\nabla f(a)$ 或 $\mathbf{grad} f(a) = \left(\dfrac{\partial f}{\partial x_i}(a)\right)_{i=1}^n \in \mathbb{R}^n$: 函数 $f: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ 在 $a \in \Omega$ 处的梯度.

$\mathbf{div}\ \boldsymbol{v}(a) = \sum_{j=1}^n \partial_j v_j(a)$: 向量值函数 $\boldsymbol{v} = (v_j): \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ 在 $a \in \Omega$ 处的散度.

$\mathbf{div}\ \boldsymbol{\sigma}(a) = \left(\sum_{j=1}^n \partial_j \sigma_{ij}(a)\right)_{i=1}^n$: 矩阵值函数 $\boldsymbol{\sigma} = (\sigma_{ij}): \Omega \subset \mathbb{R}^n \to \mathbb{M}^n$ 在 $a \in \Omega$ 处的散度.

$\mathbf{curl}\ \boldsymbol{h}(a) = (\partial_j h_i(a) - \partial_i h_j(a))_{1 \leqslant i<j \leqslant n}$: 向量值函数 $\boldsymbol{h} = (h_i): \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ 在 $a \in \Omega$ 处的旋度.

$f''(a) \in \mathcal{L}_2(X;Y)$: f 在 $a \in \Omega$ 处的二阶导数.

$\partial_{ij} f(a) = \dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) \in \mathbb{R}: f: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ 在 $a \in \Omega$ 处的二阶偏导数.

$f^{(m)}(a) \in \mathcal{L}_m(X;Y)$: 映射 f 在 $a \in \Omega$ 处的 m 阶导数.

$f^{(m)}(a)h^m = f^{(m)}(a)(h_i, h_2, \cdots, h_m) \in Y$ 当 $h_i = h$ 时, $1 \leqslant i \leqslant m$.

$\partial^{\boldsymbol{\alpha}} v(a) = \dfrac{\partial^{|\boldsymbol{\alpha}|} v}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, |\boldsymbol{\alpha}| = \alpha_1 + \cdots + \alpha_n$: 函数 $v: \Omega \subset \mathbb{R}^n \to \mathbb{R}$ 偏导数的重指标符号, 其中 $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_n) \in \mathbb{N}^n$.

同样的符号 $\partial_j f, \partial f/\partial x_j, \partial^2 f/\partial x_i \partial x_j$ 或 $\partial^{\boldsymbol{\alpha}} v$ 等也表示在分布意义下的偏导数.

向量、矩阵、张量

当把 $\mathbb{R}^n$ 中的一个向量视为矩阵时, 是把它看作列向量, 即 $n \times 1$ 矩阵.

$\boldsymbol{u}^{\mathrm{T}} = (u_1 u_2 \cdots u_n)$: 向量 $\boldsymbol{u}$ 的转置 (行向量, 即 $1 \times n$ 矩阵).

$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^{\mathrm{T}} \boldsymbol{v}$: $\mathbb{R}^n$ 中的欧氏内积.

$|\boldsymbol{u}| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}$: $\mathbb{R}^n$ 中的欧氏范数.

$\boldsymbol{u} \otimes \boldsymbol{v} = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} = (u_i v_j)$: $\mathbb{R}^n$ 中的张量积.

$\boldsymbol{u} \wedge \boldsymbol{v} = \varepsilon_{ijk} u_j v_k \boldsymbol{e}_i$: $\mathbb{R}^3$ 中的向量积, 其中定向张量 (ε_{ijk}) 是如下定义的三阶张量: $\varepsilon_{ijk} = 1$ 若 $\{i,j,k\}$ 是 $\{1,2,3\}$ 的偶置换, $\varepsilon_{ijk} = -1$ 若 $\{i,j,k\}$ 是 $\{1,2,3\}$ 的奇置换, $\varepsilon_{ijk} = 0$ 若至少有两个指标是相同的.

$\mathbb{M}^n$: 所有实 $n \times n$ 矩阵的集合.

$\mathbb{M}^{m \times n}$: 所有实 $m \times n$ 矩阵 (m 行, n 列) 的集合.

$\mathbb{U}^n$: 所有实 $n \times n$ 可逆矩阵的集合.

$\mathbb{M}_+^n = \{\boldsymbol{F} \in \mathbb{M}^n; \det \boldsymbol{F} > 0\}$.

$\mathbb{O}^n = \{\boldsymbol{P} \in \mathbb{M}^n; \boldsymbol{P}\boldsymbol{P}^{\mathrm{T}} = \boldsymbol{P}^{\mathrm{T}}\boldsymbol{P} = \boldsymbol{I}\}$: 所有实 $n \times m$ 正交矩阵的集合.

$\mathbb{O}^n_+ = \{\boldsymbol{P} \in \mathbb{O}^n; \det \boldsymbol{P} = 1\}$: 所有实 $n \times n$ 正常正交矩阵的集合.

$\mathbb{S}^n = \{\boldsymbol{B} \in \mathbb{M}^n; \boldsymbol{B} = \boldsymbol{B}^{\mathrm{T}}\}$: 所有实 $n \times n$ 对称矩阵的集合.

$S^n_>$: 所有实 $n \times n$ 对称正定矩阵的集合.

给定矩阵 $\boldsymbol{A} \in \mathbb{M}^{m\times n}$, $(\boldsymbol{A})_{ij}$ 表示其第 i 行, 第 j 列的元素.

符号 $\boldsymbol{A} = (a_{ij})$ 意指 $a_{ij} = (\boldsymbol{A})_{ij}$; 或等价地

$$\boldsymbol{A} = (a_{ij}) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

$\boldsymbol{I} = (\delta_{ij})$: 单位矩阵.

$\boldsymbol{A}^{\mathrm{T}}$: 矩阵 $\boldsymbol{A}$ 的转置.

$\boldsymbol{A}^{-1}$: 矩阵 $\boldsymbol{A}$ 的逆.

$\boldsymbol{A}^{-\mathrm{T}} = (\boldsymbol{A}^{-1})^{\mathrm{T}} = (\boldsymbol{A}^{\mathrm{T}})^{-1}$.

$\boldsymbol{A}^{1/2} \in \mathbb{S}^n_>$: 矩阵 $\boldsymbol{A} \in \mathbb{S}^n_>$ 的平方根.

$\mathbf{Diag}\ \mu_i$ 或 $\mathbf{Diag}(\mu_1, \mu_2, \cdots, \mu_n)$: 对角线元素 (依序) 为 $\mu_1, \mu_2, \cdots, \mu_n$ 的对角阵.

tr $\boldsymbol{A}$: 矩阵 $\boldsymbol{A}$ 的迹 (tr 也表示 Sobolev 空间中的迹算子).

det $\boldsymbol{A}$: 矩阵 $\boldsymbol{A}$ 的行列式.

$\lambda_i = \lambda_i(\boldsymbol{A}), 1 \leqslant i \leqslant n$: 矩阵 $\boldsymbol{A} \in \mathbb{M}^n$ 的特征值.

$|\boldsymbol{A}| = \sup_{\boldsymbol{v}\neq 0}(|\boldsymbol{A}\boldsymbol{v}|/|\boldsymbol{v}|)$: 矩阵 $\boldsymbol{A}$ 从属于欧氏范数的算子范数.

$\boldsymbol{A} : \boldsymbol{B} = \mathrm{tr}\, \boldsymbol{A}^{\mathrm{T}}\boldsymbol{B}$: $\mathbb{M}^{m\times n}$ 中的矩阵内积.

$\|\boldsymbol{A}\|_F = \{\boldsymbol{A} : \boldsymbol{A}\}^{1/2}$: 矩阵 $\boldsymbol{A} \in \mathbb{M}^{m\times n}$ 的 Frobenius 范数.

Cof $\boldsymbol{A} \in \mathbb{M}^n$: 矩阵 $\boldsymbol{A} \in \mathbb{M}^n$ 的余子式矩阵.

$\mathbb{R}^n$ 中的微分几何

拉丁下标或指数在集合 $\{1, 2, \cdots, n\}$ 中变化; 希腊下标或指数在集合 $\{1, 2\}$ 中变化; 重复的下标或指数仍沿用求和约定.

$\mathbb{E}^n$: n 维欧氏空间.

下面, Ω 是 $\mathbb{R}^n$ 的开子集, 而 $\boldsymbol{\Theta} = (\Theta_i) : \Omega \to \mathbb{R}^n$ 是足够光滑的浸入.

$\boldsymbol{g}_i = \partial_i$, $\quad g_{ij} = \boldsymbol{g}_i \cdot \boldsymbol{g}_j$, $\quad (g^{ij}) = (g_{ij})^{-1}$, $\quad \boldsymbol{g}^i = g^{ij}\boldsymbol{g}_j$.

$(g_{ij}) : \Omega \to \mathbb{S}^n_>$: 度量张量.

$\Gamma_{ijq} = \dfrac{1}{2}(\partial_j g_{iq} + \partial_i g_{jq} - \partial_q g_{ij}) = \partial_i \boldsymbol{g}_j \cdot \boldsymbol{g}_q$: 第一类 Christoffel 符号.

$\Gamma_{ij}^{p} = a^{pq}\Gamma_{ijq} = \boldsymbol{g}^p \cdot \partial_i \boldsymbol{g}_j$: 第二类 Christoffel 符号.

$v_{i\|j} = \partial_j v_i - \Gamma_{ij}^p v_p$: 向量场 $v_i \boldsymbol{g}^i : \Omega \to \mathbb{E}^n$ 的共变导数.

$R_{qijk} = \partial_j \Gamma_{ikq} - \partial_k \Gamma_{ijq} + \Gamma_{ij}^p \Gamma_{kqp} - \Gamma_{ik}^p \Gamma_{jqp}$: Riemann 曲率张量的共变分量.

下面, ω 是 $\mathbb{R}^2$ 的开子集, 而 $\boldsymbol{\theta} = (\theta_\alpha) : \omega \to \mathbb{R}^3$ 足够光滑的浸入.

$\boldsymbol{a}_\alpha = \partial_\alpha \boldsymbol{\theta}$; $\boldsymbol{a}_3 = \boldsymbol{a}^3 = \dfrac{\boldsymbol{a}_1 \wedge \boldsymbol{a}_2}{|\boldsymbol{a}_1 \wedge \boldsymbol{a}_2|}$, $\alpha_{\alpha\beta} = \boldsymbol{a}_\alpha \cdot \boldsymbol{a}_\beta$, $(a^{\alpha\beta}) = (a_{\alpha\beta})^{-1}$,

$\boldsymbol{a}^\alpha = a^{\alpha\beta}\boldsymbol{a}_\beta, b_{\alpha\beta} = \partial_\alpha \boldsymbol{a}_\beta \cdot \boldsymbol{a}_3 = -\boldsymbol{a}_\beta \cdot \partial_\alpha \boldsymbol{a}_3, b_\alpha^\beta = a^{\beta\sigma} b_{\alpha\sigma}$.

$(\alpha_{\alpha\beta}) : \omega \to \mathbb{S}_>^2$: 曲面 $\boldsymbol{\theta}(\omega)$ 的第一基本形式.

$(b_{\alpha\beta}) : \omega \to \mathbb{S}^2$: 曲面 $\boldsymbol{\theta}(\omega)$ 的第二基本形式.

$\Gamma_{\alpha\beta\tau} = \dfrac{1}{2}(\partial_\beta a_{\alpha\tau} + \partial_\alpha a_{\beta\tau} - \partial_\tau a_{\alpha\beta}) = \partial_\alpha \boldsymbol{a}_\beta \cdot \boldsymbol{a}_T$: 第一类 Christoffel 符号.

$\Gamma_{\alpha\beta}^\sigma = a^{\sigma\tau}\Gamma_{\alpha\beta\tau} = \boldsymbol{a}^\sigma \cdot \partial_\alpha \boldsymbol{a}_\beta$: 第二类 Christoffel 符号.

$\eta_{\alpha|\beta} = \partial_\beta \eta_\alpha - \Gamma_{\alpha\beta}^\sigma \eta_\sigma$: 切向量场 $\eta_\alpha \boldsymbol{a}^\alpha : \omega \to \mathbb{E}^3$ 的共变导数.

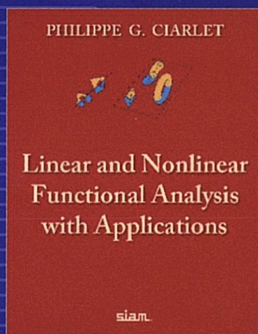
$R_{\tau\alpha\beta\sigma} = \partial_\beta \Gamma_{\alpha\sigma\tau} - \partial_\sigma \Gamma_{\alpha\beta\tau} + \Gamma_{\alpha\beta}^\mu \Gamma_{\sigma\tau\mu} - \Gamma_{\alpha\sigma}^\mu \Gamma_{\beta\tau\mu}$: 曲面 $\boldsymbol{\theta}(\omega)$ 的 Riemann 曲率张量的共变分量.

索　引

相关图书清单

序号	书号	书名	作者
1	9 787040 243086	解析函数论初步	H. 嘉当
2	9 787040 251562	微分学	H. 嘉当
3	9 787040 284171	广义函数论	L. 施瓦兹
4	9 787040 258011	微分几何——流形、曲线和曲面（修订第二版）	M. 贝尔热 等
5	9 787040 263626	拓扑学教程——拓扑空间和距离空间、数值函数、拓扑向量空间（第二版）	G. 肖盖
6	9 787040 251555	谱理论讲义（第二版）	J. 迪斯米埃
7	9 787040 246193	拟微分算子和 Nash–Moser 定理	S. 阿里纳克 等
8	9 787040 294675	解析与概率数论导引	G. 特伦鲍姆
9	9 787040 322941	概率与位势（第一卷）——可测空间	C. 德拉歇利 等
10	9 787040 319606	无穷小计算	J. 迪厄多内
11	9 787040 332384	分布系统的精确能控性、摄动和镇定（第一卷）——精确能控性	J.–L. 利翁斯
12	9 787040 287578	代数学教程	R. 戈德门特
13	9 787040 351767	完全集与三角级数	J.–P. 卡安
14	9 787040 477481	线性与非线性泛函分析及其应用（上册）	P. G. 希阿雷
15	9 787040 477498	线性与非线性泛函分析及其应用（下册）	P. G. 希阿雷
16	9 787040 495003	分析与代数原理（及数论）（第一卷）（第 2 版）	P. 科尔梅
17	9 787040 498691	分析与代数原理（及数论）（第二卷）（第 2 版）	P. 科尔梅
18	9 787040 520316	有限群导引	J.–P. 塞尔
19	9 787040 545524	线性与非线性泛函分析及其应用（上册修订版）	P. G. 希阿雷
20	9 787040 548037	线性与非线性泛函分析及其应用（下册修订版）	P. G. 希阿雷

郑重声明



这是一部涵盖线性与非线性泛函分析大部分核心课题的巨著，书中给出了基本定理及其在线性和非线性偏微分方程、以及源自于数值分析和最优化理论的专题中的各种应用。第 1 章不加证明地复述本书其他部分所需要的实分析及函数论的主要内容。第 2 到第 6 章讨论线性泛函分析及其应用。第 7、8、9 章则讨论非线性泛函分析及其应用。

本书具有如下特色：

- 它是自封闭的，对大部分定理都给出了完整的证明，其中有些不易在文献中查到，而要重构证明也有相当难度。
- 含有 400 多道习题及 50 余幅插图。
- 给出了丰富的历史注记及原始参考文献，揭示了诸多重要结果的原始思想。

本书适合本科高年级学生、研究生以及研究人员学习和参考，既可用于教学也可供读者进行自学。



学科类别：数学

academic.hep.com.cn


ISBN 978-7-04-054552-4
9 787040 545524 >

定价 89.00 元